Propensity to marry and heterogeneity in the health benefits of marriage*

Dmitry Tumin

The Ohio State University

Department of Sociology

* Direct correspondence to Dmitry Tumin, The Ohio State University, Department of Sociology, 238 Townshend Hall, 1885 Neil Avenue Mall, Columbus, OH 43210. E-mail: tumin.1@osu.edu.

Abstract:

Marriage is associated with good health, but the protective effect of marriage varies widely, such that some people experience substantial health benefits from marriage and others experience no benefit. Our study explores if the marriage effect on health is moderated by the likelihood of marrying. Using propensity score methods, we test for heterogeneity in the marriage effect on self-rated health and a scale of depressive symptoms in the National Longitudinal Survey of Youth, 1979 cohort. We find only modest differences in the marriage effect between the married and the unmarried, and no evidence that the marriage effect is positively or negatively associated with the propensity to marry. Our findings suggest that when the likelihood of marriage is defined as a composite of many early-life factors, it does not substantially moderate the health benefits of marriage, contradicting the hypothesis that the same factors discouraging marriage also make marriage less beneficial.

Fewer Americans than ever are getting married, and the role of marriage in society has changed. Cohabitation has emerged as an alternative to marriage and the proportion of children raised by single rather than married parents has increased (Casper & Bianchi 2002; Cherlin 2009; McLanahan 2004; Manning & Smock 2005). Responding to a seeming retreat from marriage, some scholars have made a case for marriage promotion (Waite 1995; Nock 2005), prominently citing the health benefits of getting and staying married (Waite & Gallagher 2000). Marriage promotion initiatives target people who are less likely to marry, and ostensibly extend the health benefits of marriage to this group (Manning, Trella, Lyons, & Du Toit 2010; Ooms & Wilson 2004; Umberson & Montez 2010). Yet it is unclear if people who do not marry would benefit from marriage (Huston & Melz 2004).

Prior studies have outlined two conflicting perspectives on this question, with each receiving some empirical support. The advantages of getting married might compensate for social and economic disadvantages that make people less likely to marry. This view is analogous to resource substitution theory, which describes greater health returns to education for women compared to men (Ross & Mirowsky 2006). The health benefits of marriage might likewise make up for social or economic disadvantages that make people both less healthy and less likely to marry. For example, people with lower incomes are less likely to marry and less likely to be in good health, but were they to marry, their marriage could be an important source of health protection that makes up for their economic disadvantage (Choi & Marks 2011). Thus, marriage might be equally or more beneficial to health among people who are unlikely to marry, provided that the effects of marriage on health—including social control of health behaviors, social support, and the sharing of economic resources (Umberson & Montez 2010)—remain significant among people who are less likely to marry.

On the other hand, the same disadvantages that reduce the probability of marriage might also lead to lower quality marriages (Lichter 2001), and a higher risk of marital dissolution (Bramlett & Mosher 2002). When a marriage is dysfunctional or short-lived, this may attenuate any health benefits that marriage might confer (Williams et al. 2011). The desire to marry is widespread, but some studies note disparities in the availability of "marriage-material" partners (Lichter 2001; Lichter, Anderson & Hayward 1995). Promoting marriage among people who want to marry but are limited in their choice of partners could increase the number of brittle or dysfunctional unions, (Huston & Melz 2004) and possibly dilute the average benefit of marriage. Returning to the above example, encouraging a low-income couple to marry may not benefit their health if their marriage is buffetted by the strains of poverty (Kroeger-D'Souza 2012).

Thus, we find two competing perspectives on whether the health benefits of marriage are positively or negatively associated with the likelihood of marrying. In this study, we apply propensity score methods to test for heterogeneity in the marriage effect on health. Our study is the first to consider the likelihood of marriage as a composite of many social and economic variables that might moderate the relationship between marriage and health outcomes. This approach allows us to test for heterogeneity in the marriage effect across a composite measure of propensity to marry, rather than across a few separate components of this propensity. Consequently, our analysis answers two broad questions implied by the debate over marriage promotion: first, whether unmarried adults could expect a comparable health benefit to marrying as those adults who have already married; and second, whether adults experience different health benefits of marriage depending on their propensity to have married.

Health benefits of marriage

Marriage is linked to better mental and physical health (Averett et al. 2012; Hughes & Waite 2009; Wood, Goesling, & Avellar 2007) and lower mortality risk (Dupre, Beck & Meadows 2009). Marriage can improve health through multiple pathways, including social support, social control of health behaviors, and access to resources that can be used to protect health (Umberson & Montez 2010; Williams & Umberson 2004). Although cross-sectional differences in health between married and unmarried adults are profound (Waite & Gallagher 2000), studies investigating health changes after marriage have found inconsistent evidence that the transition to marriage improves health (Williams & Umberson 2004). For example, Averett and colleagues (2008) found entry into marriage is associated with an increased risk of obesity, and Wu and Hart (2002) found that entering a marriage was not associated with improvements in either mental or physical health over a two-year period. These findings contradict studies that show health to improve after marriage (Wood et al. 2007), suggesting marriage may benefit health only modestly or inconsistently.

Why might marriage have a modest effect on health? First, the health benefits of marriage appear to vary across outcomes: some outcomes, such as physical health, may take longer to evince the benefits of marriage (Hughes & Waite 2009), and others, such as body weight, may be adversely affected by marriage (Averett et al. 2012). Studies using a short follow-up period may not detect all the health benefits marriage eventually confers (Wu & Hart 2002). Second, selection on prior health status—whether positive selection, whereby healthier people are more likely to marry; or adverse selection, whereby less healthy people are more likely to marry—may confound the estimate of health benefits ascribed to marrying (Fu & Goldman 1996; Lillard & Panis 1996). An estimate of the average marriage effect must take into

account selection into marriage, or *the pre-marriage heterogeneity bias* (Wood et al. 2007). Recent studies have examined multiple health outcomes and have used fixed effects methods to deal with some of the problems of selection into marriage (Averett et al. 2012; Musick & Bumpass 2012). These studies show marriage has small beneficial effects for a variety of health outcomes; but adverse effects on other health outcomes, and, in the case of mental health, an effect that dissipates relatively quickly.

Third, the modest size of the average effect ascribed to marriage may conceal substantial variation in this effect: marriage may be very beneficial for some people, but it may have no effect on others, or even an adverse effect for some. We refer to this variation as *heterogeneity in the marriage effect on health*, and elaborate on it in the following section. If such heterogeneity is present, then the average marriage effect represents a weighted average of heterogeneous effects. Among these effects, some will be higher and others will be lower than the average, and some may even be in the opposite direction of the average effect (Brand & Xie 2010). Therefore, estimates of the average marriage effect must be supplemented with a systematic study of variation across individuals in the extent to which marriage improves health (Morgan & Winship 2007). Studies that have used fixed-effects estimators have attempted to mitigate the pre-marriage heterogeneity bias, but have not yet addressed the marriage effect heterogeneity bias. By testing for significant heterogeneity in the marriage effect, we take a further step towards accurately characterizing the relationship between marriage and health (Morgan & Todd 2008).

Heterogeneity in the marriage effect

Prior studies have found many dimensions of heterogeneity in the marriage effect. The health benefits of marriage appear to vary across demographic groups. The protective effect of marriage

may also be contingent on economic circumstances, prior childbearing, the quality of the current marriage, initial health status, and the outcome of the marriage—if it endures or ends. As an example of heterogeneity in the marriage effect, among Black women, but not White women, marriage *increases* the risk of health limitations, and among men, marriage reduces the risk of health limitations more strongly for Whites as compared to Blacks (Teachman 2010a). Furthermore, the link between marriage and lower substance use is also much weaker among Black young adults than White or Hispanic young adults (Kroeger-D'Souza 2012). Although these findings suggest social disadvantage attenuates the benefits of marriage, other studies reach the opposite conclusion. For instance, among men, lower incomes amplify the longevity advantage associated with being married rather than never married (Choi & Marks 2011).

Other moderators of the marriage effect include premarital fertility and the quality of the marriage. When single mothers marry, they do not experience the same protective effect as childless women who marry (Williams et al. 2008). Among White and Hispanic women who have a premarital birth, an enduring marriage to the child's father may still predict better health at midlife, compared to remaining single (Williams et al 2011). Some marriages may also be unhappy, or of low quality. Unhappy marriages do not confer the same mental health benefits as happy marriages (Hawkins & Booth 2005; Frech & Williams 2007). On the other hand, both high- and low-conflict marriages appear to increase longevity among low-income men (Choi & Marks 2011).

Health status is another important moderator of the protective effect of marriage. Indirectly, poor health or emergent health problems may diminish marital quality, leading to marital dissolution (Joung et al. 1998), and negating the health benefits of having married in the first place. For example, the emergence of health problems limiting one's ability to work or the

kind of work one could do predicts an increased risk of divorce among men (Teachman 2010b). Zheng and Thomas (2013) also propose that low self-rated health may directly attenuate the longevity advantage of being married. Marriage may be more effective in preventing health problems from emerging, rather than compensating for pre-existing poor health. Among people who have existing health problems, a spouse's support for healthy behaviors may do little to help them recover from their condition or disease (Zheng & Thomas 2013). Conversely, Frech and Williams (2007) find that marriage improves mental health (measured by a scale of depressive symptoms) to a greater extent among people who were depressed before marrying than among people who were not depressed before marrying.

Finally, many health benefits of marrying are lost if that marriage ends (Dupre & Meadows 2007; Hughes & Waite 2009). Remarriages appear to be less beneficial for both mental and physical health, compared to first marriages (Barrett 2000; Hughes & Waite 2009). At midlife, currently married people who have only married once are healthier than currently married people who have experienced the break-up of a previous marriage (Hughes & Waite 2009). The accumulation of years in the married state may amplify the health advantage of married people (Dupre et al. 2009; Dupre & Meadows 2007), while the accumulation of marital break-ups and years spent divorced and widowed may undo the health advantages of marrying (Dupre & Meadows 2007; Hughes & Waite 2009). These findings show that marriage is not simply a one-time intervention that could improve health, but a process that unfolds over the life course, with multiple possible outcomes (remaining married, divorcing, and so on).

Does the marriage effect vary by differential selection into marriage?

Socioeconomic status, race/ethnicity, premarital childbearing, and health status may influence a

person's likelihood of marrying and the likelihood of that marriage enduring (Bramlett & Mosher 2002; Fu & Goldman 1996; Qian, Lichter & Mellot 2005). These variables not only influence selection into marriage, but also appear to moderate the protective effect of marriage, as discussed above. Consequently, scholars have questioned whether people who are unlikely to marry would reap the health benefits of marriage if they do marry (Huston & Melz 2004; Williams et al. 2011). In other words, the marriage effect on health may vary systematically with the likelihood of marriage. This proposition challenges one of the tenets of the marriage promotion movement, which holds that the benefits of marriage are similar between people who are likely to marry and those who are unlikely to marry (Huston & Melz 2004; Lichter 2001).

Prior studies have considered how specific variables affecting the likelihood of marrying might moderate the health effect of marriage. Some evidence suggests people who are unlikely to marry are also unlikely to benefit from marriage. For example, single mothers have poorer marriage prospects than childless women (Qian, Lichter & Mellot 2005), and tend to benefit less from marriage if they do marry (Williams et al. 2008; Williams et al. 2011). Other studies, however, challenge this conclusion. Choi and Marks (2011) find marriage to have greater benefits for longevity among low-income men compared to high-income men, even though the former are substantially less likely to get married (Sweeney 2002). Furthermore, studies of subpopulations and communities in which marriage is rare find that continuous marriage nevertheless improves health behaviors and mental health (Ali & Ajilore 2011; Green et al. 2012; Hill et al. 2012).

Although past studies have investigated specific predictors of marriage as potential moderators of the marriage effect on health, they have not sought to test whether the health benefits of marriage are moderated by a general measure of the probability of marriage. Rather,

these studies have examined how the effect of marriage varies across several variables—such as race/ethnicity and premarital fertility—that are known to affect the likelihood of marriage. One limitation of this approach is that only a few potential moderators may be examined at one time. Therefore, heterogeneity in the marriage effect observed across levels of one variable (e.g., premarital fertility) might in fact be due to differences in another variable (e.g., growing up in a single-parent household) not included in the model as a potential moderator of the marriage effect. Viewed in a counterfactual framework, prior studies on heterogeneity in the marriage effect may be comparing married and unmarried groups that have dissimilar characteristics. The observed heterogeneous marriage effect may be caused not by the variable of interest, but by other differences in the characteristics of the married and the unmarried groups. There are many group differences in socioeconomic status, relationship and fertility history, culture and attitudes, and other observed or unobserved factors that could plausibly moderate the health effect of marriage. In order to distill the heterogeneous marriage effect attributable to a single variable of interest, one would need to control for interaction terms among marriage and every one of its known predictors, as well as their combinations—an impractical proposition.

The ignorability assumption requires that there be no omitted variables simultaneously affecting both selection into marriage and the outcome variable. Under this assumption, the only interaction consequential for selection bias (both pre-marriage selection bias and marriage effect heterogeneity bias) is between marriage and the propensity of marriage (Xie et al. 2012; Heckman et al. 2006). Although the ignorability assumption may be violated in observational data, using propensity score matching under this assumption has advanced our understanding of the marriage effect (Williams et al. 2011). In this study, we will control for a wide range of confounders that affect both the likelihood of marriage and its effects on health, thereby making

the ignorability assumption more plausible. Unlike previous studies that have used propensity score matching to estimate the average marriage effect, we pursue two ways of assessing whether the marriage effect on health varies by differential selection into marriage, under the assumption of ignorability.

Our first approach towards assessing heterogeneity in the marriage effect is to compare the average treatment (marriage) effect for the treated (married) to the average treatment effect for the untreated (not married). This method relies on calculating a propensity score for the transition to marriage, based on all the available variables known to influence the chances of this transition (Ali & Ajilore 2011; Williams et al. 2011). This approach tests whether people who did not marry would benefit as much from marriage as people who did marry. Morgan (2001) applies this method to study whether the "Catholic school effect" on academic achievement would extend to students who do not attend Catholic schools. A second, related approach is to examine how the health effects of marriage vary across levels of the propensity score. This approach investigates whether people who were least likely to marry benefited as much as those who were most likely to marry. Brand and Xie (2010) apply this method to study heterogeneous returns to a college education. In this study, we use both methods to test for heterogeneity in the health effects of marriage, addressing the long-standing debate on whether promoting marriage can serve to promote health.

Data and measures

Data

We use data from the National Longitudinal Study of Youth, 1979 cohort (NLSY79) (BLS 2010). The NLSY79 was begun with a cohort of 12,686 Americans who were between 14 and 22

years old in 1979. Respondents were interviewed annually until 1994, and biannually since then. We use data from the baseline 1979 interview as well as the 1998, 2000, 2002, 2004 and 2006 waves, during which the 40+ health module was administered to respondents. This module collected extensive health measures from respondents who had reached age 40, with each respondent completing the module once between 1998 and 2006, depending on their birth year (CHRR 2012).

Of the original 12,686 respondents, we exclude 2,923 respondents in the military and economically disadvantaged non-Hispanic White oversamples, which were mostly discontinued after 1986, and therefore did not complete the 40+ health module. We then exclude 971 respondents who had ever married before the 1979 interview. Of the remaining 8,792 respondents, we exclude 1,471 cases missing data on either outcome measure (self-rated health or depressive symptoms scale, described in detail below), with this missingness mainly due to attrition from the sample before the 40+ module was administered. This exclusion means our results may be biased by selective attrition from the survey (e.g., single people being more difficult to track down for reinterview), but, due to high response rates to the 40+ health module among respondents interviewed in 1998-2006, our results are unlikely to be substantially biased by selective non-response to the health questions. Finally, we exclude cases missing data on baseline covariates used to construct the propensity score, yielding an analytic sample of 3,231 men and 3,149 women[1]. To adjust for unequal probabilities of selection into the sample and attrition, we apply custom frequency weights to all our analyses (CHRR 2012). All analyses were performed using Stata/MP 13.0 (College Station, TX: StataCorp).

[1] This sample size applies to a parsimonious specification of the propensity score model. When additional baseline covariates are included, the size of the analytic sample decreases.

Health outcomes

We use two measures from the 40+ health module to capture the status of respondents' physical and mental health. The first measure is the respondent's rating of their general health on a five-point scale: 5, "excellent;" 4, "very good;" 3, "good;" 2, "fair;" or 1, "poor." Self-rated health is highly correlated with morbidity and mortality risk (Idler, Russell & Davis, 2000; Idler & Benyamini, 1997), and is commonly used in studies assessing the health effects of marriage (Williams & Umberson 2004; Hughes & Waite 2009; Williams et al. 2011). Our second measure is a truncated version of the Center for Epidemiological Studies Depression (CES-D) Scale (CHRR 2012). The full CES-D scale contained 20 items from previously validated scales and its validity was assessed with respect to clinical evaluations of depression and other self-report measures (Radloff 1977). In the 40+ health module, respondents were presented with a subset of seven items from the CES-D scale describing symptoms of depression, and were asked how often they had experienced each over the past week.

Symptoms covered by the CES-D scale included: "I did not feel like eating; my appetite was poor;" "I had trouble keeping my mind on what I was doing;" "I felt depressed;" "I felt that everything I did was an effort;" "my sleep was restless;" "I felt sad;" and "I could not get `going'" (CHRR 2012). For each item, responses were coded as follows: 0, rarely or none of the time; 1, some or a little of the time; 2, occasionally or a moderate amount of the time; and 3, most or all of the time. Scores were summed over all seven items to yield the score on the overall scale, the range of which is 0-21 points. Cases in which one or more of the items in the CES-D scale were missing were treated as missing on the summary CES-D variable, and were omitted from the sample. We treat both the self-rated health and the CES-D outcomes as continuous

variables, with higher values on the self-rated health scale indicating better health and higher values on the CES-D scale indicating worse health.

Marital history

We consider marriage as a binary "treatment" which may or may not improve health (Williams et al. 2011). Yet the distinction between "married" and "unmarried" can be drawn in different ways. One possibility is simply dividing respondents by their marital status at a single point in time, distinguishing between currently married and currently unmarried. This approach overlooks variation in marital histories: some currently married adults could have been divorced in the past, and, given that remarriage tends to be associated with weaker health benefits than first marriage (Dupre et al. 2009; Hughes & Waite 2009), this may lead to understating the health benefits of getting and staying married. As prior studies find continuous marriage to be most consistently associated with health benefits, we also consider a second method of defining the marriage treatment: distinguishing between those who get and stay married and all others. Here, the reference group includes the never married, the divorced, separated or widowed, and the remarried.

People do not know at the time of marriage if their marriage will last; and indeed, many marriages do not (Huston & Melz 2004; Lichter 2001). Therefore, estimates of the marriage effect based on enduring marriages could be upwardly biased relative to the benefits newlyweds should expect to reap. Hence, we consider a third, less restrictive measures of marriage, which compares respondents who have ever married to respondents who have never married, regardless of whether these marriages last. This measure of ever marrying represents the worst-case

estimate, as it collapses married, divorced, and remarried respondents in the same "married" category, despite prior findings that divorce largely undoes the health benefits of marriage.

Estimates of the marriage effect may be sensitive not only to the definition of marriage, but also to when marriage is measured. Measurements taken too close to when the outcomes are assessed risk understating the marriage effect, as some health outcomes, especially measures of physical health, take time to respond to changes in marital status (Hughes & Waite 2009). On the other hand, measurements taken too far from the outcomes (i.e., too early in the life course), do not account for any marital transitions that may follow. For example, if the ever-married measure is assessed too early in the life course, many people in the "never married" group will, in fact, marry before the outcome measures are assessed, diluting the contrast between the two groups. In the main analysis, we measure marriage as of age 40. As most respondents complete the 40+ health module within two years of turning 40, this means that our measure of marriage precedes the measures of the outcome variables by 0 to 2 years. A few respondents miss the next wave of the NLSY after turning 40, and for them, the measures of marriage and the measures of the outcome variables are more distant from one another—up to but not more than ten years apart. We then discuss the consequences of measuring marriage at age 38 or age 35 instead.

We use the NLSY79's constructed marital history variables to calculate three measures of marriage. NLSY79 staff combine data on marital status at each interview with data on recent marital transitions to produce a set of start dates for the first, second, and third marriage and a set of end dates for the first and second marriage (CHRR 2012; Haurin 1994). As very few respondents ever exit a third marriage, we consider any respondent who enters a third marriage (4% of men and 5% of women in the analytic sample) to be married from that point onwards. The measure of current marriage is coded 1 if the respondent married or remarried before age 40,

and if that marriage did not end or ended after the respondent turned 40; and 0 otherwise. The measure of being continuously married is coded as 1 if the respondent has ever married by age 40 and the end date of their first marriage is either later than their 40th birthday or missing; and 0 otherwise. The measure of ever marrying is coded as 1 if the respondent had ever married by age 40, and as 0 otherwise.

Statistical Models

Propensity score of marriage

For each definition of marriage, we aim to estimate its effect on the two health outcomes using propensity score methods (Caliendo & Kopeinig, 2008; Rosenbaum & Rubin 1983). Briefly, propensity score methods seek to draw causal inference from observational data under the assumption of ignorability. Let *Y* be an outcome variable, and *D* assignment to a binary treatment (i.e., cases are divided into treated and untreated). If *X* is a vector of covariates determining assignment to the treated (*D = 1*) or control (*D* = 0) group, the propensity score is simply the probability of being in the treated group conditional on the variables in *X*. The ignorability assumption states that being in the treated or control group will be independent of the outcome of interest, conditional on the propensity score.

$$(Y^1, Y^0) \perp D \mid P(D = 1|X)$$

where ⊥ denotes independence, and *P(D = 1|X)* is the propensity score.

The choice of variables included in the propensity score model is key to obtaining valid estimates. Ideally, variables included in the propensity score model should be ones that confound the association between the treatment and the outcome—in our case, between marriage and health (Caliendo & Kopeinig 2008; Williams et al. 2011). As the propensity score approach requires covariates to be assessed prior to the "treatment" (i.e., marriage), we do not include

adulthood socioeconomic characteristics (e.g., income, educational attainment, or employment status) in the propensity score models. The NLSY79 cohort is 14-22 years old at the 1979 baseline, and marriage as we define it may happen any time from that point until respondents reach age 40. Some indicators, such as educational attainment, are inappropriate for the age range of respondents at baseline, and including indicators measured after 1979 risks violating the causal order between the covariates and the treatment—for instance, one's income in 1989 may be partially determined by one's marital history over the preceding decade.

One approach to specifying the propensity score model is to select a parsimonious set of covariates that are clearly expected to affect both the likelihood of the treatment and the outcome (Caliendo & Kopeinig 2008). Another approach calls for overspecifying the model, taking advantage of as many variables as can be plausibly included (Millimet & Tchernis 2009). We estimate the propensity score using both approaches; in the main analysis, with a parsimonious model that satisfies the balancing property, and in supplemental analyses, with an exhaustive model that includes additional variables and extensive interactions among them, but does not satisfy the balancing property. We calculate the propensity scores separately for the men's and women's subsamples. This approach leads us to slightly different specifications of the parsimonious model for men and women. In Table 1, we summarize the specifications of our parsimonious and exhaustive propensity score models and the construction of the variables used in these models.

[Table 1 about here]

Heterogeneous treatment effects

Our primary interest in applying propensity score methods is to examine heterogeneity in treatment effect, or, in this case, the effect of marriage on health. A key type of heterogeneity in the treatment effect is the difference between the treatment effect for the treated (ATT) and the treatment effect for the controls (ATC), or those who do not receive the treatment (Morgan 2001; Morgan & Todd 2008). At the individual level, we denote the effect of the treatment (marriage) δ and write an ordinary least squares bivariate regression where the outcome *Y* is a function of *D*, denoting membership in the treated or control group:

$$Y_i = \alpha + \delta D_i + \varepsilon_i$$

Let d_i represent individual membership in the treated ($d_i = 1$) or control ($d_i = 0$) group, and $\hat{p}_i$ represent the individual's estimated propensity score, or probability of being in the treated group. We wish to weight the regression equation to obtain two estimates of the treatment effect: one for the average treatment effect on the treated, and another for the average treatment effect on the controls. With w_i representing the individual's weight, we calculate the following weights to obtain the ATT:

$$\text{For } d_i = 1\text{: } w_i = 1$$

$$\text{For } d_i = 0\text{: } w_i = \frac{\hat{p}_i}{1-\hat{p}_i}$$

To obtain the ATC instead, we calculate the following weights:

$$\text{For } d_i = 1\text{: } w_i = \frac{1-\hat{p}_i}{\hat{p}_i}$$

$$\text{For } d_i = 0\text{: } w_i = 1$$

Now, the regression equation weighted by either the ATT or ATC weights will return an estimate of the treatment effect δ that represents the ATT or ATC, respectively. The average treatment effect, or ATE, can be similarly computed by using the following weights:

$$\text{For } d_i = 1\text{: } w_i = \frac{1}{\hat{p}_i}$$

$$\text{For } d_i = 0\text{: } w_i = \frac{1}{1-\hat{p}_i}$$

Morgan and Todd (2008) suggest two criteria for judging whether this procedure reveals causal effect heterogeneity. The first criterion is the substantive magnitude of the difference between ATT and ATC. The second criterion is overlap in the 95% confidence intervals of the ATT and ATC estimates, although the presence of some overlap does not necessarily indicate that the null hypothesis (ATT = ATC) should be rejected. A formal statistical test of the difference between ATT and ATC derived from this method is, unfortunately, not available.

A second aspect of heterogeneity in the causal effect is the difference in the treatment effect among people with differential probabilities of receiving the treatment (Brand & Xie 2010; Xie et al. 2012). In our case, this represents the difference between people who were most likely to marry and those who were least likely to marry, as indicated by their respective propensity scores. The method we use recasts the individual-level estimated effect of the treatment as a function of the propensity score stratum. The result is the following hierarchical linear model:

$$Y_{ij} = \alpha_j + \delta_j D_{ij} + \varepsilon_{ij}$$

$$\delta_j = \delta_0 + \phi j + \eta_j$$

where *i* indexes the case, *j* indexes the propensity score stratum, ϕ is the linear slope of the treatment effect across propensity score strata, and the level 2 error, η_j, is assumed to be normal. The linear slope of the treatment effect is the main parameter of interest here. For example, suppose the effect of marriage (*D*) on the outcome (*Y*) is generally positive, as might be the case if *Y* represented a summary measure of good health. Then, a positive slope, $\phi > 0$, would indicate that people who are more likely to marry benefit more from marriage, and people who are less likely to marry benefit less from marriage. If the relationship between treatment *D* and

outcome Y were generally negative, we would draw the same conclusion if the trend in the treatment effect across propensity strata were negative, or if $\phi < 0$.

Results

Sample characteristics

Table 2 summarizes pre-marriage covariates for men and women in the NLSY79 that are used to calculate the parsimonious propensity score. We divide respondents into two categories according to their marital status at age 40: married and unmarried. (Descriptive statistics for all covariates, and additional comparisons of descriptive statistics, including between ever married and never married and between continuously married and not continuously married, are presented in Appendix I.) The weighted proportions shown in Table 2 illustrate how men and women who are married at age 40 differ at the 1979 baseline from their peers who are unmarried at age 40. For example, the unmarried are more likely to be non-Hispanic Black, more likely to have grown up in poverty, and are less likely to have expected to marry by age 24 when they were first interviewed. Among women, only 3% of those who are married at age 40 had children in 1979, compared to 7% of those who were unmarried at age 40. For a few variables (e.g., urban residence in 1979), the contrast between the married and unmarried subsamples is less apparent. Nevertheless, including these covariates in the propensity score model helps balance the means of other covariates between treated (married) and untreated (unmarried) men or women.

[Table 2 about here]

Average health effects of marriage

Table 3 summarizes the proportions of men and women coded as receiving the "treatment" according to each definition of marriage, and the average treatment effects (ATE) for the

outcomes of self-rated health and the CES-D scale of depressive symptoms. An overwhelming majority of men (80%) and women (84%) in the sample had married by the time they reached age 40. At age 40, 68% of men and 69% of women were married, and 52% of men and 51% of women were still in their first marriage. Relative to their unmarried peers, both men and women who were married at age 40 had a statistically significant advantage in self-rated health (ATE = 0.17) and on the CES-D scale of depressive symptoms (ATE = -0.90 and -0.95, respectively). The signs on these estimated effects are in the expected directions, indicating the married have better self-rated health and fewer depressive symptoms than the unmarried. The effect of marriage on depressive symptoms is smaller when marriage is defined as having ever married by age 40. Under this definition, the ATE is -0.51 among men and -0.53 among women, or about 55% of the ATE when defining marriage as being currently married. On the other hand, the average effect of marriage on self-rated health among men is similar whether marriage is defined as being currently married or ever married. Among women, the average marriage effect on self-rated health is slightly greater when comparing the ever-married to the never-married (ATE = 0.21) than when comparing currently married to the currently unmarried.

[Table 3 about here]

Effect heterogeneity between the married and the unmarried

Next, we consider how the benefits of marriage differ between the married and the unmarried. In Table 4, we report results from weighted least squares regressions contrasting the average health effect of marriage on the married (ATT) with the average health effect of marriage on the unmarried (ATC). All of the effects are in the same direction as the ATEs in Table 3: marriage is associated with fewer depressive symptoms (negative effect) and higher self-rated health (positive effect). Therefore, we will discuss only the difference in magnitude

between ATT and ATC for each case. The effect of marriage on the CES-D score is stronger among the married (ATT) than the unmarried (ATC) among both men and women, with the difference between ATT and ATC being 0.06 and 0.04, respectively. The differences in the marriage effect on CES-D score between ATT and ATC are comparable when defining marriage as being continuously married at age 40. However, when marriage is defined as having ever married by age 40, ATC is greater than ATT among both men and women. The latter difference indicates that ever marrying would affects CES-D score more strongly among people who never marry than people who have married as of age 40. In the case of self-rated health, we do not find any differences between ATT and ATC greater than 0.01 among women, regardless of the definition of marriage. Among men, the ATT of marriage on self-rated health exceeds ATC by 0.01 when marriage is defined as being currently or continuously married. When marriage is defined as having ever married by age 40, ATC exceeds ATT among men by 0.02.

[Table 4 about here]

Our comparisons of ATT and ATC show that the effect of marriage on the married is larger than the effect of marriage on the unmarried when marriage is defined as current or continuous marriage, but smaller when marriage is defined as having ever married. The differences between ATT and ATC, however, are very small in the case of self-rated health. Unfortunately, there is no statistical test for the difference in treatment effects; following Morgan and Todd (2008), we compare the magnitudes and the confidence intervals of estimated effects instead. Considering CES-D scores, the greatest difference between ATC and ATT among women is 0.15, using the "ever married" definition of the treatment. This difference is slightly less than a third of the smaller of the two effect estimates, ATT = -0.47. Furthermore, the confidence intervals of the ATC and ATT estimates overlap by 80%, using the width of the

confidence interval for the ATT as the reference. Among men, the ATC-ATT difference is somewhat larger (0.24). This difference is about half of the smaller effect estimate, ATT = -0.45, but the ATT and ATC confidence intervals still overlap by 58%. Considering self-rated health scores, the ATT and ATC are equal (rounding to the nearest .01) using every definition of the treatment among women. The greatest difference among men is 0.02, using the "ever married" definition. This difference equals about 13% of the smaller effect estimate, ATT = 0.15, and the ATT and ATC confidence intervals overlap by 84%. In short, interpreting the ATT-ATC difference for which the strongest case can be made, never-married men would benefit more from marriage than ever-married men by only a quarter of a point on a 21-point scale of depressive symptoms. This difference, however, is smaller and in the opposite direction if the treatment is defined as continuous or current marriage.

Effect heterogeneity across differential propensities of marriage

Does the effect of marriage vary by the likelihood of having married? We apply a hierarchical linear model to estimate the linear trend in the effect of marriage across propensity strata (Brand & Xie 2010). This trend represents the linear relationship between propensity to marry and the health advantage of marriage. If people who are more likely to marry also benefited more from marriage, the trend in the marriage effect across propensity strata would be negative in the case of the CES-D scale (i.e., people who are more likely to marry would experience a greater reduction in depressive symptoms), and positive in the case of self-rated health (i.e., people who are more likely to marry would experience a greater gain to self-rated health). Table 5 summarizes our estimates of linear heterogeneity in the marriage effect. In the case of self-rated health, we find no evidence of significant heterogeneity in the marriage effect

for either men or women, regardless of how marriage is defined (Sample-specific treatment effects are presented in Appendix II.) In the case of CES-D score, we find no significant heterogeneity in the effect of either current or continuous marriage among both men and women.

[Table 5 about here]

The only evidence of heterogeneity in the marriage effect on CES-D score among men is the slope of marriage effects across propensity strata when marriage is defined as having ever married. This slope (0.46) is positive and statistically significant, suggesting that men who are more likely to ever marry experience a smaller decrease in depressive symptoms if they do marry, compared to men who are less likely to marry. We examine this result more closely in Figure 1, which plots the treatment effects within each stratum with their 95% confidence intervals, as well as the linear trend. In the middle three propensity strata, there is no significant effect of ever marrying on CES-D score. In the lowest stratum (men least likely to ever marry), there is a significant negative effect, indicating fewer depressive symptoms due to marriage. In the highest stratum (men most likely to ever marry), there is a significant positive effect, indicating more depressive symptoms due to marriage. The latter finding is anomalous: the direction of this effect is contrary to the ATE, ATT, and ATC estimates reported in Tables 2 and 3, and we fail to replicate this finding using other definitions of marriage, such as being currently or continuously married.

[Figure 1 about here]

Sensitivity analysis

So far, we have considered the effect of marital status as of age 40 on self-rated health and depressive symptoms assessed in the next survey wave, typically within the next two years.

We redefine the treatment variables (definitions of marriage) to represent marital status as of ages 38 or 35, and repeat the above analysis, presenting the results in Appendix III. We find that the estimates of ATE are consistent with our main analysis when the outcome is self-rated health, and are somewhat smaller when the outcome is CES-D score. Differences between the marriage effect on the married and unmarried (ATT and ATC, respectively) are sometimes higher, and sometimes lower, when marital status is assessed at an earlier age than 40. For example, in the case of the CES-D score, the largest differences between ATT and ATC at age 35 are 0.11, for the effect of ever marrying among women, and 0.20, for the effect of ever marrying among men. In both cases, the 95% confidence intervals exhibit substantial overlap: 85% and 66%, respectively, relative to the width of the confidence interval for the ATT estimate. Furthermore, estimates of the linear trend in the treatment effect across propensity strata are similar in magnitude to results from the main analysis (Table 5) and are not statistically significant, except in the case of the ever-married effect on men's CES-D score.

Our analyses have so far relied on a propensity score calculated using a limited number of covariates and interaction terms (Table 1). We now turn to an exhaustive model of the propensity score, including additional covariates and interaction terms. The results are presented in Appendix IV. Using the exhaustive specification, our propensity score models no longer satisfy the balancing property, meaning that some covariates have significantly different means between married and unmarried people in the same stratum of propensity to marry. An exhaustive specification of the propensity score leads to estimates of the average marriage effect (ATE) that are similar to the main analysis. For example, the effect of being currently married on CES-D score among men is -0.90 using the parsimonious propensity score model, and -0.80 using the exhaustive model. Some differences between ATT and ATC estimates are somewhat

larger than in the main analysis. For example, the ATC of ever marrying by 40 on CES-D score is 0.42 points greater than the ATT among women, as compared to a difference of 0.15 using the parsimonious propensity score model. The overlap in the 95% confidence intervals in this case is 49.7% of the width of the 95% confidence interval for the ATT. Among men, the ATC of ever marrying on the CES-D score exceeds the ATT by 0.58 points, compared to 0.24 points in the main analysis. Here, the confidence intervals overlap by 0.1%. Although the differences between ATT and ATC are larger, linear trends in the marriage effect across propensity strata remain small and statistically insignificant when we use the exhaustive propensity score model.

Discussion and conclusions

The health benefits of marriage are known to vary. Our study tests for two types of systematic heterogeneity in the marriage effect on health: the difference between the married and unmarried; and the difference between those who are more likely to have married and those who are less likely to have married. We find weak support for the first kind of effect heterogeneity among both men and women. Using a scale of depressive symptoms as a measure of mental health, the effect of marriage on those who have ever married (ATT) is somewhat smaller than the effect of marriage estimated for those who have never married (ATC). But, when a different definition of marriage is used (i.e., contrasting between continuously married respondents and all others, or between currently married and currently unmarried), the difference is in the opposite direction, with the effect of marriage slightly stronger among the married than the unmarried. No matter what measure of marriage is used, the confidence intervals between ATT and ATC exhibit substantial overlap. For a five-point scale of general physical health, the ATT-ATC difference in the marriage effect is even more neglibile, and the overlap in their confidence

intervals is more pronounced. We find almost no evidence for the second kind of effect heterogeneity: for almost every combination of gender, definition of marriage, and health outcome, the marriage effect neither increases nor decreases across different levels of propensity to marry. In short, we find no consistent relationship between the likelihood of marriage and the effect of marriage on either self-rated health or depressive symptoms.

Our findings contrast sharply with prior studies reporting heterogeneity in the marriage effect. For example, Choi and Marks (2011) find that low-income men, who are unlikely to get married and stay married, receive a greater longevity benefit from continuous marriage than other men. Their study suggests health benefits of continuous marriage are stronger among men who are less likely to be married. Conversely, Kroeger-D'Souza (2012) finds that low socioeconomic status reduces the effectiveness of marriage in promoting healthy behaviors. As low socioeconomic status is negatively correlated with the chances of marrying (Sweeney 2002), this suggests the propensity to marry should be negatively correlated with the benefit of marriage. We infer the same hypothesis from Williams and colleagues' (2008) finding that the self-rated health benefits of continuous marriage are attenuated for single mothers, the latter being less likely to marry than childless women. Our findings offer no support to either position, however. Among both men and women, the slopes of the marriage effect across levels of marriage propensity fail to reach statistical significance in all cases but one (Table 5).

The novelty of our findings is attributable to the advantages of the model we apply in this study. A person's chances of marriage may vary for many reasons, and the propensity score approach allows us to use a composite measure of the likelihood of marriage to investigate heterogeneity in the marriage effect. Prior studies consider heterogeneity in this effect along only a few dimensions (typically just one dimension) at a time. Although the propensity score method

cannot rule out bias due to unobserved confounders, it is well suited to simultaneously handling many sources of effect heterogeneity. Under the assumption of ignorability (i.e., ruling out unobserved confounders), the only interaction consequential for selection bias (both pre-marriage selection bias and marriage effect heterogeneity bias) is between marriage and the propensity of selection into marriage (Xie et al. 2012; Heckman et al. 2006). Our results, obtained using propensity score methods, offer weaker support for heterogeneity in the marriage effect than past studies examining interactions between marriage and a few select moderating variables. This may be because our analysis captures a wider range of confounding effects on the link between marriage and health, and these cancel out one another.

By studying heterogeneous treatment effects across levels of the propensity to marry, we compare married respondents to unmarried respondents who have similar social, demographic, and economic characteristics. By analogy, children from disadvantaged families are more likely to attend a lower-quality school, but an analysis of heterogeneity in the school effect reveals that they benefit more from school than children from advantaged families, when school quality is held constant (Brand & Xie 2010). Our paper applies this rationale to the marriage effect. For example, racial/ethnic and socioeconomic differences in marital stability (Bramlett & Mosher 2002) may explain corresponding differences in the health effects of marriage. Green and colleagues (2012) show that among urban African Americans, continuous marriage leads to healthier behaviors, including less smoking and drug use. Similarly, continuous marriage predicts lower levels of distress among low-income mothers (Hill, Reid & Reczek 2012). As continuous marriage appears to be protective even in groups where marriage is rare, we argue that differences in the average marriage effect would be attenuated when we adjust for predictors of marital disruption. The modest and mixed evidence we find for heterogeneity in the marriage

effect using propensity score methods is consistent with this argument. More generally, observed differences in the benefits of marriage across groups could be due to differences in the characteristics of the average marriage in each group, and comparable marriages could have comparable benefits even across racial/ethnic, socioeconomic, or other social divisions.

The data and methods of our study pose some limitations that restrict the generalizability of our results. First, our data come from a single cohort; contemporary and future policies will invariably deal with younger cohorts, in which the likelihood of marriage has become more polarized (McLanahan 2004). Second, we may be using an incomplete model of the propensity score, biasing our results. Our conclusions, however, are robust to a choice between an exhaustive and parsimonious specification of this model, but other relevant variables may have been omitted. Third, we consider only two health outcomes; the marriage effect could be more heterogeneous in the case of mortality (Choi & Marks 2011) or distress (Williams et al. 2008). On the other hand, the CES-D and self-rated health scales have been previously used as general indicators of the marriage effect on health (Averett et al. 2012; Hughes & Waite 2009). Though the effect of marriage may be more heterogeneous for one outcome than another, we see no theoretical justification why the effect should only be heterogeneous for some other, unexamined outcome and practically homogenous for the two outcomes we have selected.

Despite these limitations, our study extends an active line of inquiry into the health consequences of marriage. Recent studies demonstrate ongoing interest in accurately characterizing the average effect of marriage (Averett et al. 2012; Musick & Bumpass 2012) and identifying groups for which it is more or less salient (Williams et al. 2011). One argument for marriage promotion policy, that the protective health effects of marriage are large and consistent across people, reoccurs as a touchstone in these studies. We take on this argument directly, and

find that the effect of marriage is modest, but surprisingly homogenous between people who are more likely to marry and people who are less likely to marry. Marriage improves self-rated health, measured on a five-point scale, by at most a quarter of a point. Marriage improves mental health, as measured by a 21-point scale of depressive symptoms, by about one point, at most. Other research confirms the average benefits of marriage are modest and not consistent across all measures of health (Averett et al. 2012; Musick & Bumpass 2012).

The debate over marriage promotion policies has, in part, been a debate over the wisdom of getting people to marry who otherwise would remain single. At its core, the marriage promotion movement proposes that extending the institution of marriage beyond those people who select into it might have some benefits, including benefits to health (Waite & Gallagher 2000). Yet people could be unlikely to marry and unlikely to benefit from marriage for similar reasons: economic hardship, racial discrimination, the lack of "marriage material" potential partners, and premarital childbearing, among others. Some studies have supported this view, while others have found evidence that people who are less likely to marry for a particular reason (e.g., race/ethnicity or income level) can nevertheless benefit from marriage, sometimes more so than people who are more likely to marry. Extending the conclusions of the latter studies, our findings fail to refute the rationale for marriage promotion. People who are unlikely to marry, according to our analysis, do not *all* share characteristics that blunt the health benefits of marriage. So marriage promotion, despite a checkered record of efficacy, may, after all, make a contribution to public health—if the appropriate population is targeted and appropriate, effective interventions are used. These conditions pose a challenge that is formidable, but, in principle, surmountable.

References

Ali, M. M., & Ajilore, O. (2011). Can marriage reduce risky health behavior for African-Americans? *Journal of Family and Economic Issues*, 32(2), 191-203.

Averett, S. L., Argys, L. M., & Sorkin, J. (2012). In sickness and in health: an examination of relationship status and health using data from the Canadian National Public Health Survey. *Review of Economics of the Household*.

Averett, S. L., Sikora, A., & Argys, L. M. (2008). For better or worse: relationship status and body mass index. *Economics and Human Biology, 6*(3), 330-349.

Barrett, A. E. (2000). Marital trajectories and mental health. *Journal of Health and Social Behavior, 41*(4), 451-464.

Bramlett, M. D., & Mosher, W. D. (2002). Cohabitation, marriage, divorce and remarriage in the United States. *Vital Health Statistics*, 23 (22). Hyattsville, MD: National Center for Health Statistics.

Brand, J. E., & Xie, Y. (2010). Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education. *American Sociological Review*, 75(2), 273-302.

Bureau of Labor Statistics (BLS), U. S. Department of Labor. (2010). *National Longitudinal Survey of Youth 1979 cohort, 1979-2010 (rounds 1-24) [computer file]*.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31-72.

Casper, L. M., & Bianchi, S. M. (2002). *Continuity and Change in the American Family*. Thousand Oaks, CA: Sage Publications.

Center for Human Resource Research (CHRR), the Ohio State University. (2012). National Longitudinal Survey of Youth 1979. Retrieved 01/24/2012, from http://www.nlsinfo.org/content/cohorts/NLSY79

Cherlin, A. J. (2009). *The Marriage-Go-Round: The State of Marriage and the Family in America Today*. New York: Vintage Books.

Choi, H., & Marks, N. F. (2011). Socioeconomic status, marital status continuity and change, marital conflict, and mortality. *Journal of Aging and Health*, 23(4), 714-742.

Dupre, M. E., Beck, A. N., & Meadows, S. O. (2009). Marital trajectories and mortality among US adults. *American Journal of Epidemiology*, 170(5), 546-555.

Dupre, M. E., & Meadows, S. O. (2007). Disaggregating the effects of marital trajectories on health. *Journal of Family Issues*, 28(5), 623-652.

Frech, A., & Williams, K. (2007). Depression and the psychological benefits of entering marriage. *Journal of Health and Social Behavior*, 48(2), 149-163.

Fu, H., & N. Goldman. (1996). Incorporating health into models of marriage choice: demographic and sociological perspectives. *Journal of Marriage and the Family*, 58(3), 740-758.

Green, K. M., Doherty, E. E., Fothergill, K. E., & Ensminger, M. E. (2012). Marriage trajectories and health risk behaviors throughout adulthood among urban African Americans. *Journal of Family Issues*, 33(12), 1595-1618.

Haurin, R. J. (1994). *Marriage and Childbearing of Adults: an Evaluation of the 1992 National Longitudinal Survey of Youth*. Columbus, OH: Center for Human Resource Research, The Ohio State University.

Hawkins, D. N., & Booth, A. (2005). Unhappily ever after: effects of long-term, low-quality marriages on well-being. *Social Forces, 84*(1), 451-471.

Hill, T. D., Reid, M., & Reczek, C. (2012). Marriage and the mental health of low-income urban women with children. *Journal of Family Issues*, forthcoming.

Hughes, M. E., & Waite, L. J. (2009). Marital biography and health at mid-life. *Journal of Health and Social Behavior*, 50, 344-358.

Huston, T. L., & Melz, H. (2004). The case for (promoting) marriage: the devil is in the details. *Journal of Marriage and Family, 66*, 943-958.

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of Health and Social Behavior*. 38(March), 21-37.

Idler, E. L., Russell, L.B. & Davis, D. (2000). Survival, functional limitations, and self-rated health in the NHANES I Epidemiologic Follow-up Study, 1992. *American Journal of Epidemiology*. 152(9), 874-883.

Joung, I. M. A., van de Mheen, H. D., Stronks, K., van Poppel, F. W. A., & Mackenbach, J. P. (1998). A longitudinal study of health selection in marital transitions. *Social Science & Medicine, 46*(3), 425-435.

Kroeger-D'Souza, R. A. (2012). *Union formation, parenthood, and health risk indicators among the socially disadvantaged*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. 3529908).

Lichter, D. T. (2001). Marriage as public policy. Washington, DC: Progressive Policy Institute.

Lichter, D. T., Anderson R. N., & Hayward, M. D. (1995). Marriage markets and marital choice. *Journal of Family Issues*, 16(4), 412-431.

Lillard, L. A., & Panis, C. W. A. (1996). Marital status and mortality: the role of health. *Demography, 33*(3), 313-327.

Manning, W. D., Trella, D., Lyons, H., & Du Toit, N. C. (2010). Marriageable women: a focus on participants in a community healthy marriage program. *Family Relations*, 59(1), 87-102.

Manning, W. D., & Smock, P. J. (2005). Measuring and modeling cohabitation: new perspectives from qualitative data. *Journal of Marriage and Family, 67*(4), 989-1002.

McLanahan, S. (2004). Diverging destinies: how children are faring under the second demographic transition. *Demography, 41*(4), 607-627.

Millimet, D. L., & Tchernis, R. (2009). On the specification of propensity scores, with applications to the analysis of trade policies. *Journal of Business & Economic Statistics*, 27(3), 397-415.

Morgan, S. L. (2001). Counterfactuals, causal effect heterogeneity, and the Catholic school effect on learning. *Sociology of Education*, 74, 341-374.

Morgan, S. L., & Todd, J. J. (2008). A diagnostic routine for the detection of consequential heterogeneity of causal effects. *Sociological Methodology*, 38(1), 231-281.

Morgan, S. L., & Winship, C. (2007). *Counterfactuals and Causal Inference: Methods & Principles for Social Research.* New York: Cambridge University Press.

Musick, K. and L. Bumpass. (2012). Reexamining the case for marriage: union formation and changes in well-being. *Journal of Marriage and Family*, 74(1), 1-18.

Nock, S. (2005). Marriage as a public issue. *The Future of Children*, 15(2), 13-32.

Ooms, T. J., & Wilson, P. C. (2004). The challenges of offering couples and marriage education to low income couples. *Family Relations*, 53, 440-446.

Qian, Z., Lichter, D. T., & Mellott, L. M. (2005). Out-of-wedlock childbearing, marital prospects and mate selection. *Social Forces,* 84(1), 473-491.

Radloff, L. S. (1977). The CES-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement,* 1(3), 385-401.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika,* 70(1), 41-55.

Ross, C. E., & Mirowsky, J. (2006). Sex differences in the effect of education on depression: resource multiplication or resource substitution? *Social Science & Medicine*, 63, 1400-1413.

Ross, C. E., & Mirowsky, J. (2010). Gender and the health benefits of education. *The Sociological Quarterly*, 51, 1-19.

Sweeney, M. M. (2002). Two decades of family change: the shifting economic foundations of marriage. *American Sociological Review, 67,* 132-147.

Teachman, J. (2010a). Family life course statuses and transitions: relationships with health limitations. *Sociological Perspectives*, 53(2), 201-220.

Teachman, J. (2010b). Work-related health limitations, education, and the risk of marital disruption. *Journal of Marriage and Family*, *72*(4), 919-932.

Umberson, D., & Montez, J. K. (2010). Social relationships and health: a flashpoint for health policy. *Journal of Health and Social Behavior*, 51(1), S54-S56.

Waite, L. J. (1995). Does marriage matter? *Demography, 32*, 483-507.

Waite, L. J., & Gallagher, M. (2000). *The Case for Marriage: Why Married People are Happier, Healthier, and Better Off Financially*. New York: Doubleday.

Williams, K., Sassler, S., Frech, A., Addo, F., & Cooksey, E. (2011). Nonmarital childbearing, union history, and women's health at midlife. *American Sociological Review*, 76(3), 465-486.

Williams, K., S. Sassler, and L. M. Nicholson. (2008). For better or for worse? The consequences of marriage and cohabitation for single mothers. *Social Forces*, 86(4):1481-1511.

Williams, K., & Umberson, D. (2004). Marital status, marital transitions, and health: a gendered life course perspective. *Journal of Health and Social Behavior, 45*(1), 81-98.

Wood, R. G., Goesling, B., & Avellar, S. (2007). The effect of marriage on health: a synthesis of recent research evidence. Washington, DC: US Department of Health and Human Services.

Wu, Z., & Hart, R. (2002). The effects of marital and nonmarital union transitions on health. *Journal of Marriage and Family, 64*(2), 420-432.

Xie, Y., Brand, J. E., & Jann, B. (2012). Estimating heterogeneous treatment effects with observational data. *Sociological Methodology*, 42, 314-347.

Zheng, H., & Thomas, P. A. (2013). Marital status, self-rated health, and mortality: overestimation of health or diminishing protection of marriage? *Journal of Health and Social Behavior*, 54(1), 128-143

Table 1. Variables used to construct propensity score models and their inclusion in parsimonious and exhaustive propensity score models

Name	Description	Type	Parsimonious model		Exhaustive model
			Women	Men	
Mother's education	Years of schooling completed	Continuous	no	no	yes
Father's education	Years of schooling completed	Continuous	no	no	yes
Intact family at 14	Lived with both biological parents at age 14	Binary	yes	yes	yes
Siblings	Number of living siblings	Continuous	no	no	yes
Urban residence	Urban rather than rural residence	Binary	yes	no	yes
Poverty status	Family income below poverty line	Binary	yes	yes	yes
Health limitation	Health limits amount or type of work	Binary	no	no	yes
Any children	Respondent lives with any of their children	Binary	yes	yes	yes
Religious upbringing	Catholic, Baptist, other Protestant, or other	Categorical	no	yes	yes
Expected age at marriage	Respondent expects to marry by age 24	Binary	yes	yes	yes
Expect marriage in five years	Respondent expects to be married in five years	Binary	no	yes	yes
Race/ethnicity	Non-Hispanic White, Non-Hispanic Black, or Hispanic	Categorical	yes*	yes*	yes**
Birth date	Date of birth in century-month code	Continuous	no	no	yes**

* Includes interactions with all other variables in the model

** Includes interactions with all other variables in the model and birth date-race/ethnicity three-way interactions

Table 2. Weighted proportions of pre-marriage characteristics, by gender and marital status

	Women (n = 3,149)				Men (n = 3,231)			
	Married at age 40		Unmarried at age 40		Married at age 40		Unmarried at age 40	
	b	(SE)	b	(SE)	b	(SE)	b	(SE)
Race/ethnicity								
Non-Hispanic White	0.83	(0.01)	0.65	(0.02)	0.84	(0.01)	0.69	(0.01)
Non-Hispanic Black	0.11	(0.01)	0.28	(0.01)	0.11	(0.01)	0.24	(0.01)
Hispanic	0.06	(0.003)	0.07	(0.01)	0.06	(0.003)	0.07	(0.01)
Lived in intact family at 14	0.79	(0.01)	0.67	(0.02)	0.79	(0.01)	0.69	(0.02)
Urban residence at 14	0.80	(0.01)	0.79	(0.02)	n/a		n/a	
Lived in poverty in 1979	0.11	(0.01)	0.23	(0.01)	0.10	(0.01)	0.18	(0.01)
Any children in 1979	0.03	(0.004)	0.07	(0.01)	0.01	(0.002)	0.01	(0.003)
Religious upbringing								
Catholic	n/a		n/a		0.21	(0.01)	0.27	(0.02)
Baptist	n/a		n/a		0.29	(0.01)	0.27	(0.02)
Other Protestant	n/a		n/a		0.34	(0.01)	0.30	(0.02)
All others	n/a		n/a		0.16	(0.01)	0.16	(0.01)
Expect to marry by age 24	0.68	(0.01)	0.55	(0.02)	0.49	(0.01)	0.40	(0.02)
Expect to marry in five years	n/a		n/a		0.44	(0.01)	0.34	(0.02)

n/a : Covariate not used to calculate parsimonious model of the propensity score. See text and Table 1 for details.

Table 3. Proportion married and average health effects of marriage, by gender and definition of marriage.

	Women (n = 3,149)					Men (n = 3,231)				
		ATE (SE)					ATE (SE)			
Definition of marriage	Prop-ortion married	CES-D scale		Self-rated health		Prop-ortion married	CES-D scale		Self-rated health	
Currently married	0.69	-0.95***	(0.17)	0.17 ***	(0.04)	0.68	-0.90***	(0.15)	0.17***	(0.04)
Contin-uously married	0.51	-0.97***	(0.16)	0.16***	(0.04)	0.52	-0.78***	(0.14)	0.15***	(0.04)
Ever married	0.84	-0.53*	(0.22)	0.21***	(0.05)	0.80	-0.51**	(0.17)	0.16**	(0.05)

* $p < .05$; ** $p < .01$; *** $p < .001$, two-tailed tests.

Table 4. Comparison of marriage effects between married (ATT) and unmarried (ATC)

	Women (n = 3,149)				Men (n = 3,231)			
	b (SE)				b (SE)			
Definition of marriage	CES-D scale		Self-rated health		CES-D scale		Self-rated health	
Currently married								
ATT weights	-0.97***	(0.18)	0.17***	(0.04)	-0.92***	(0.16)	0.17***	(0.04)
ATC weights	-0.93***	(0.19)	0.17***	(0.04)	-0.86***	(0.16)	0.16***	(0.04)
Continuously married								
ATT weights	-1.01***	(0.16)	0.16***	(0.04)	-0.79***	(0.14)	0.15***	(0.04)
ATC weights	-0.95***	(0.17)	0.16***	(0.04)	-0.76***	(0.14)	0.14***	(0.04)
Ever married								
ATT weights	-0.47	(0.24)	0.22***	(0.06)	-0.45*	(0.18)	0.15**	(0.05)
ATC weights	-0.62**	(0.23)	0.22***	(0.05)	-0.69*	(0.18)	0.17***	(0.04)

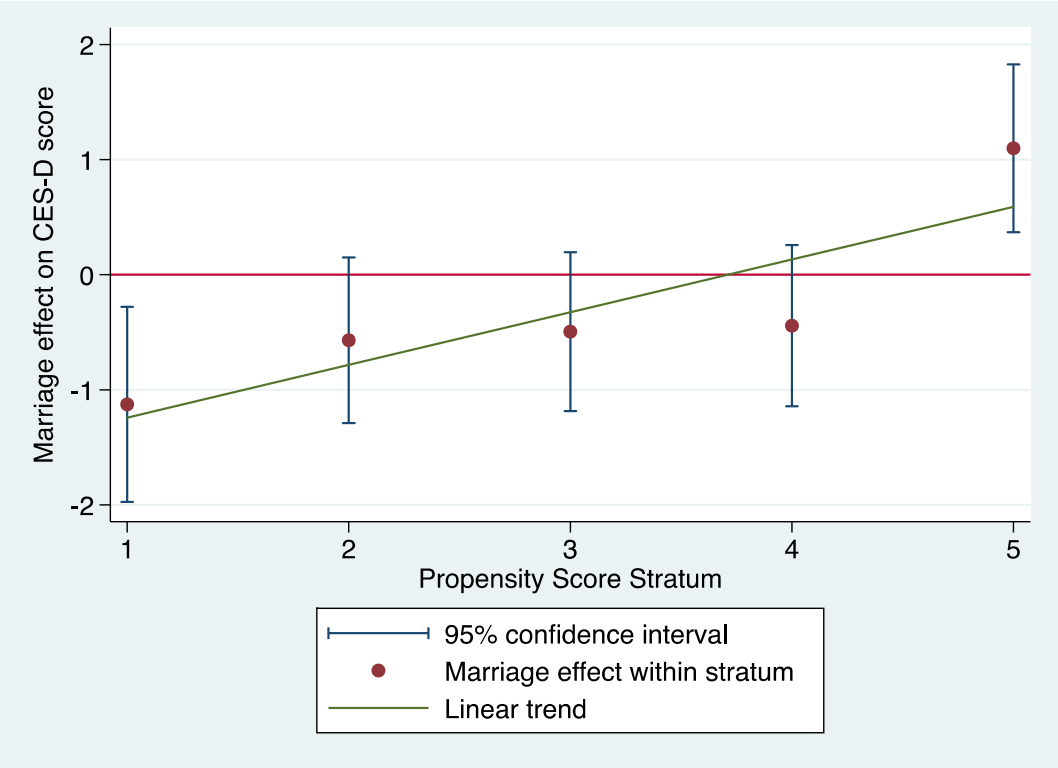
* $p < .05$; ** $p < .01$; *** $p < .001$, two-tailed tests.

Table 5. Linear relationship between stratum of marriage propensity and estimate of marriage effect

	Women (*n* = 3,149)				Men (*n* = 3,231)			
	b (SE)				b (SE)			
Definition of marriage	CES-D scale		Self-rated health		CES-D scale		Self-rated health	
Currently married	-0.03	(0.10)	0.01	(0.02)	0.00	(0.11)	0.01	(0.03)
Continuously married	-0.12	(0.17)	0.01	(0.04)	0.03	(0.12)	0.01	(0.03)
Ever married	0.09	(0.16)	0.00	(0.04)	0.46***	(0.12)	-0.04	(0.04)

*** $p < .001$, two-tailed tests.

Figure 1. Effect of ever marrying by age 40 on CES-D score among men, by propensity score stratum.



Appendix I. Complete descriptive statistics for analytic sample.

Weighted proportions or means of pre-marriage characteristics, by propensity score model specification and marital status at age 40: NLSY79 Women

Parsimonious model			Currently married at age 40				Continuously married as of age 40				Ever married as of age 40			
(*n* = 3,149)	All		No		Yes		No		Yes		No		Yes	
	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)
Race/ethnicity														
Non-Hispanic White	0.77	(0.01)	0.65	(0.01)	0.83	(0.01)	0.72	(0.01)	0.82	(0.01)	0.54	(0.02)	0.82	(0.01)
Non-Hispanic Black	0.17	(0.01)	0.28	(0.01)	0.11	(0.01)	0.21	(0.01)	0.12	(0.01)	0.39	(0.02)	0.12	(0.01)
Hispanic	0.06	(0.01)	0.07	(0.01)	0.06	(0.003)	0.07	(0.004)	0.06	(0.004)	0.07	(0.01)	0.06	(0.003)
Lived in intact family at 14	0.75	(0.01)	0.67	(0.02)	0.79	(0.01)	0.69	(0.01)	0.81	(0.01)	0.66	(0.02)	0.77	(0.01)
Urban residence at 14	0.80	(0.01)	0.79	(0.02)	0.80	(0.01)	0.79	(0.01)	0.81	(0.01)	0.80	(0.02)	0.80	(0.01)
Lived in poverty in 1979	0.15	(0.01)	0.23	(0.01)	0.11	(0.01)	0.18	(0.01)	0.11	(0.01)	0.27	(0.02)	0.12	(0.01)
Any children in 1979	0.04	(0.003)	0.07	(0.01)	0.03	(0.004)	0.06	(0.01)	0.03	(0.004)	0.10	(0.01)	0.03	(0.003)
Expect to marry by age 24	0.64	(0.01)	0.55	(0.02)	0.68	(0.01)	0.60	(0.01)	0.68	(0.01)	0.47	(0.02)	0.68	(0.01)

Exhaustive model			Currently married at age 40				Continuously married as of age 40				Ever married as of age 40			
(*n* = 2.602)	All		No		Yes		No		Yes		No		Yes	
	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)
Race/ethnicity														
Non-Hispanic White	0.78	(0.01)	0.65	(0.02)	0.83	(0.01)	0.72	(0.01)	0.83	(0.01)	0.54	(0.02)	0.82	(0.01)
Non-Hispanic Black	0.17	(0.01)	0.29	(0.01)	0.11	(0.01)	0.21	(0.01)	0.12	(0.01)	0.40	(0.02)	0.12	(0.01)
Hispanic	0.06	(0.003)	0.06	(0.01)	0.06	(0.003)	0.06	(0.01)	0.05	(0.01)	0.06	(0.01)	0.06	(0.003)
Lived in intact family at 14	0.75	(0.01)	0.67	(0.02)	0.79	(0.01)	0.70	(0.01)	0.81	(0.01)	0.66	(0.02)	0.77	(0.01)
Number of siblings	3.30	(0.04)	3.50	(0.07)	3.21	(0.05)	3.36	(0.06)	3.25	(0.06)	3.58	(0.10)	3.25	(0.05)
Urban residence at 14	0.81	(0.01)	0.80	(0.02)	0.81	(0.01)	0.80	(0.01)	0.81	(0.01)	0.81	(0.02)	0.81	(0.01)
Lived in poverty in 1979	0.14	(0.01)	0.22	(0.01)	0.11	(0.01)	0.18	(0.01)	0.11	(0.01)	0.26	(0.02)	0.12	(0.01)
Health limitation in 1979	0.04	(0.01)	0.04	(0.01)	0.04	(0.01)	0.04	(0.01)	0.05	(0.01)	0.05	(0.01)	0.04	(0.01)
Any children in 1979	0.04	(0.003)	0.07	(0.01)	0.03	(0.004)	0.06	(0.01)	0.03	(0.004)	0.10	(0.01)	0.03	(0.003)
Religious upbringing														
Catholic	0.25	(0.01)	0.32	(0.02)	0.22	(0.01)	0.30	(0.01)	0.21	(0.01)	0.34	(0.02)	0.23	(0.01)
Baptist	0.27	(0.01)	0.25	(0.02)	0.28	(0.01)	0.26	(0.01)	0.28	(0.01)	0.24	(0.02)	0.28	(0.01)
Other protestant	0.33	(0.01)	0.26	(0.02)	0.36	(0.01)	0.28	(0.01)	0.38	(0.01)	0.23	(0.02)	0.35	(0.01)
All others	0.15	(0.01)	0.18	(0.01)	0.14	(0.01)	0.17	(0.01)	0.14	(0.01)	0.19	(0.02)	0.15	(0.01)
Expect to marry by age 24	0.64	(0.01)	0.55	(0.02)	0.68	(0.01)	0.60	(0.01)	0.68	(0.01)	0.46	(0.02)	0.68	(0.01)
Expect to marry in 5 years	0.54	(0.01)	0.47	(0.02)	0.57	(0.01)	0.50	(0.01)	0.57	(0.02)	0.43	(0.02)	0.56	(0.01)
Date of birth (CMC)	736.22	(0.57)	735.35	(0.96)	736.63	(0.70)	738.10	(0.77)	734.43	(0.82)	731.85	(1.30)	737.08	(0.62)

Weighted proportions or means of pre-marriage characteristics, by propensity score model specification and marital status at age 40: NLSY79 Men

Parsimonious model			Currently married at age 40				Continuously married as of age 40				Ever married as of age 40			
(*n* = 3,231)	All		No		Yes		No		Yes		No		Yes	
	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)
Race/ethnicity														
Non-Hispanic White	0.79	(0.01)	0.69	(0.01)	0.84	(0.01)	0.74	(0.01)	0.84	(0.01)	0.62	(0.02)	0.83	(0.01)
Non-Hispanic Black	0.15	(0.01)	0.24	(0.01)	0.11	(0.01)	0.20	(0.01)	0.10	(0.01)	0.30	(0.02)	0.11	(0.01)
Hispanic	0.06	(0.003)	0.07	(0.01)	0.06	(0.003)	0.07	(0.004)	0.06	(0.004)	0.08	(0.01)	0.06	(0.003)
Lived in intact family at 14	0.76	(0.01)	0.69	(0.02)	0.79	(0.01)	0.71	(0.01)	0.80	(0.01)	0.67	(0.02)	0.78	(0.01)
Lived in poverty in 1979	0.13	(0.01)	0.18	(0.01)	0.10	(0.01)	0.16	(0.01)	0.10	(0.01)	0.22	(0.02)	0.11	(0.01)
Any children in 1979	0.01	(0.002)	0.01	(0.003)	0.01	(0.002)	0.01	(0.002)	0.01	(0.003)	0.01	(0.01)	0.01	(0.002)
Religious upbringing														
Catholic	0.23	(0.01)	0.27	(0.02)	0.21	(0.01)	0.28	(0.01)	0.19	(0.01)	0.27	(0.02)	0.22	(0.01)
Baptist	0.28	(0.01)	0.27	(0.02)	0.29	(0.01)	0.26	(0.01)	0.30	(0.01)	0.27	(0.02)	0.29	(0.01)
Other protestant	0.33	(0.01)	0.30	(0.02)	0.34	(0.01)	0.29	(0.01)	0.36	(0.01)	0.31	(0.02)	0.33	(0.01)
All others	0.16	(0.01)	0.16	(0.01)	0.16	(0.01)	0.17	(0.01)	0.15	(0.01)	0.15	(0.02)	0.16	(0.01)
Expect to marry by age 24	0.46	(0.01)	0.40	(0.02)	0.49	(0.01)	0.45	(0.02)	0.15	(0.01)	0.35	(0.02)	0.49	(0.01)
Expect to marry in 5 years	0.41	(0.01)	0.34	(0.02)	0.44	(0.01)	0.38	(0.02)	0.47	(0.02)	0.30	(0.02)	0.43	(0.01)
Exhaustive model			Currently married at age 40				Continuously married as of age 40				Ever married as of age 40			
(*n* = 2,684)	All		No		Yes		No		Yes		No		Yes	
	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)
Race/ethnicity														
Non-Hispanic White	0.79	(0.01)	0.69	(0.01)	0.84	(0.01)	0.74	(0.01)	0.84	(0.01)	0.62	(0.02)	0.83	(0.01)
Non-Hispanic Black	0.15	(0.01)	0.24	(0.01)	0.10	(0.01)	0.20	(0.01)	0.10	(0.01)	0.30	(0.02)	0.11	(0.01)
Hispanic	0.06	(0.003)	0.07	(0.01)	0.06	(0.004)	0.07	(0.004)	0.06	(0.004)	0.08	(0.01)	0.06	(0.003)
Lived in intact family at 14	0.76	(0.01)	0.69	(0.02)	0.79	(0.01)	0.71	(0.01)	0.81	(0.01)	0.67	(0.02)	0.78	(0.01)
Number of siblings	3.22	(0.04)	3.41	(0.07)	3.13	(0.05)	3.30	(0.06)	3.15	(0.06)	3.45	(0.09)	3.16	(0.05)
Urban residence at 14	0.80	(0.01)	0.81	(0.01)	0.79	(0.01)	0.79	(0.01)	0.81	(0.01)	0.83	(0.02)	0.79	(0.01)
Lived in poverty in 1979	0.13	(0.01)	0.18	(0.01)	0.10	(0.01)	0.16	(0.01)	0.10	(0.01)	0.21	(0.02)	0.10	(0.01)
Health limitation in 1979	0.04	(0.004)	0.05	(0.01)	0.04	(0.01)	0.05	(0.01)	0.04	(0.01)	0.05	(0.01)	0.04	(0.01)
Any children in 1979	0.01	(0.002)	0.01	(0.003)	0.01	(0.002)	0.01	(0.002)	0.01	(0.003)	0.01	(0.01)	0.01	(0.002)
Religious upbringing														
Catholic	0.23	(0.01)	0.27	(0.02)	0.21	(0.01)	0.28	(0.01)	0.19	(0.01)	0.27	(0.02)	0.22	(0.01)
Baptist	0.28	(0.01)	0.27	(0.02)	0.29	(0.01)	0.26	(0.01)	0.31	(0.01)	0.26	(0.02)	0.29	(0.01)
Other protestant	0.33	(0.01)	0.30	(0.02)	0.34	(0.01)	0.29	(0.01)	0.36	(0.02)	0.31	(0.02)	0.33	(0.01)
All others	0.16	(0.01)	0.16	(0.01)	0.16	(0.01)	0.17	(0.01)	0.15	(0.01)	0.15	(0.02)	0.16	(0.01)
Expect to marry by age 24	0.46	(0.01)	0.40	(0.02)	0.49	(0.01)	0.45	(0.02)	0.47	(0.02)	0.34	(0.02)	0.49	(0.01)
Expect to marry in 5 years	0.41	(0.01)	0.34	(0.02)	0.44	(0.01)	0.38	(0.02)	0.43	(0.02)	0.29	(0.02)	0.43	(0.01)
Date of birth (CMC)	734.5	(0.58)	735.8	(0.98)	733.9	(0.72)	736.9	(0.82)	732.3	(0.81)	736.0	(1.20)	734.2	(0.66)

Appendix II. Detailed HTE results.

Propensity score stratum ranges and stratum-specific treatment effects

	Currently married					Continuously married					Ever married				
Women (n = 3,149)	Score range	Treatment effect: CES-D scale		Treatment effect: self-rated health		Score range	Treatment effect: CES-D scale		Treatment effect: self-rated health		Score range	Treatment effect: CES-D scale		Treatment effect: self-rated health	
		b	(SE)	b	(SE)		b	(SE)	b	(SE)		b	(SE)	b	(SE)
Stratum 1	[0, 0.4)	-0.27	(0.62)	0.14	(0.15)	[0, 0.4)	-0.72*	(0.36)	0.26**	(0.08)	[0,0.6)	-0.27	(0.50)	0.26*	(0.11)
Stratum 2	[0.4, 0.5)	-1.18**	(0.41)	0.17	(0.09)	[0.4, 0.5)	-1.46***	(0.35)	0.12	(0.09)	[0.6, 0.7)	-1.33**	(0.50)	0.24*	(0.09)
Stratum 3	[0.5, 0.6)	-0.29	(0.67)	-0.14	(0.12)	[0.5, 0.6)	-1.43***	(0.35)	0.18*	(0.08)	[0.7, 0.8)	-0.85	(0.61)	-0.05	(0.13)
Stratum 4	[0.6, 0.7)	-1.39**	(0.47)	0.31**	(0.11)	[0.6, 1.0]	-0.93**	(0.34)	0.23**	(0.08)	[0.8, 0.9)	-0.57	(0.46)	0.18	(0.10)
Stratum 5	[0.7, 0.75)	-1.03**	(0.38)	0.17*	(0.08)	-	-		-		[0.9, 1.0]	-0.07	(0.54)	0.32**	(0.12)
Stratum 6	[0.75, 1.0]	-0.83	(0.45)	0.21*	(0.10)	-	-		-		-	-		-	
Men (n = 3,231)															
Stratum 1	[0, 0.4)	-0.35	(1.04)	0.23	(0.30)	[0, 0.3)	-0.82	(0.72)	-0.03	(0.17)	[0, 0.6)	-1.13**	(0.43)	0.26*	(0.10)
Stratum 2	[0.4, 0.5)	-1.32**	(0.39)	0.18	(0.09)	[0.3, 0.35)	-0.73	(0.59)	0.22	(0.16)	[0.6, 0.7)	-0.57	(0.37)	0.12	(0.09)
Stratum 3	[0.5, 0.6)	-0.82*	(0.41)	0.18	(0.10)	[0.35, 0.4)	-0.94	(0.49)	0.17	(0.13)	[0.7, 0.8)	-0.50	(0.35)	0.23*	(0.10)
Stratum 4	[0.6, 0.7)	-0.51*	(0.26)	0.18	(0.07)	[0.4, 0.5)	-1.09***	(0.29)	0.25**	(0.08)	[0.8, 0.9)	-0.44	(0.36)	0.08	(0.10)
Stratum 5	[0.7, 0.75)	-1.29*	(0.50)	0.23*	(0.12)	[0.5, 0.6)	-0.79**	(0.23)	0.18**	(0.06)	[0.9,1.0]	1.10**	(0.37)	0.04	(0.19)
Stratum 6	[0.75, 0.8)	-0.56	(0.51)	0.11	(0.13)	[0.6, 1.0]	-0.76	(0.40)	0.12	(0.11)	-	-		-	
Stratum 7	[0.8, 1.0)	-1.61*	(0.65)	0.43	(0.23)	-	-		-		-	-		-	

* $p < .05$; ** $p < .01$; *** $p < .001$, two-tailed tests.

Strata with fewer than 10 cases were joined to the next higher or lower stratum if they were the first or last, respectively.

Appendix III. Marital status or history assessed at ages 35 and 38.

Proportion married and treatment effect estimates by gender, age and definition of marriage

Definition of marriage	Prop. marr-ied	ATE				ATT				ATC				HTE linear slope			
		CES-D scale		Self-rated health		CES-D scale		Self-rated health		CES-D scale		Self-rated health		CES-D scale		Self-rated health	
		b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)
Women (n = 3,149) *As of age 35*																	
Currently married	0.67	*Results not reported as parsimonious model of propensity score fails to converge*															
Continuously married	0.53	-0.74***	(0.16)	0.16***	(0.04)	-0.76***	(0.15)	0.16***	(0.04)	-0.75***	(0.17)	0.17***	(0.04)	0.06	(0.15)	-0.03	(0.03)
Ever married	0.80	-0.41*	(0.20)	0.20***	(0.05)	-0.38	(0.21)	0.20***	(0.05)	-0.50*	(0.20)	0.19***	(0.04)	0.10	(0.11)	-0.02	(0.03)
As of age 38																	
Currently married	0.68	-0.77***	(0.17)	0.16***	(0.04)	-0.77***	(0.18)	0.17***	(0.04)	-0.79***	(0.18)	0.14***	(0.04)	-0.04	(0.11)	0.01	(0.03)
Continuously married	0.52	-0.80***	(0.16)	0.15***	(0.04)	-0.82***	(0.15)	0.15***	(0.04)	-0.82***	(0.16)	0.15***	(0.04)	-0.06	(0.15)	-0.02	(0.03)
Ever married	0.82	-0.43*	(0.21)	0.18***	(0.05)	-0.40	(0.22)	0.19***	(0.05)	-0.56**	(0.21)	0.17***	(0.05)	0.12	(0.12)	-0.01	(0.03)
Men (n = 3,231) *As of age 35*																	
Currently married	0.65	-0.64***	(0.15)	0.17***	(0.04)	-0.62***	(0.15)	0.16***	(0.04)	-0.76***	(0.15)	0.18***	(0.04)	0.12	(0.10)	-0.01	(0.03)
Continuously married	0.54	-0.51***	(0.14)	0.16***	(0.04)	-0.56***	(0.13)	0.16***	(0.04)	-0.61***	(0.15)	0.17***	(0.04)	0.11	(0.17)	-0.03	(0.04)
Ever married	0.76	-0.40*	(0.16)	0.14**	(0.04)	-0.33*	(0.16)	0.12**	(0.05)	-0.62***	(0.16)	0.16***	(0.04)	0.56***	(0.12)	-0.03	(0.04)
As of age 38																	
Currently married	0.67	-0.69***	(0.15)	0.19***	(0.04)	-0.70***	(0.15)	0.19***	(0.04)	-0.76***	(0.15)	0.18***	(0.04)	0.05	(0.11)	0.03	(0.03)
Continuously married	0.53	-0.65***	(0.14)	0.18***	(0.04)	-0.69***	(0.13)	0.18***	(0.04)	-0.70***	(0.14)	0.18***	(0.04)	0.08	(0.16)	-0.001	(0.04)
Ever married	0.79	-0.50**	(0.17)	0.17***	(0.04)	-0.43*	(0.17)	0.16**	(0.05)	-0.72***	(0.17)	0.19***	(0.04)	0.28**	(0.10)	-0.01	(0.03)

* $p < .05$; ** $p < .01$; *** $p < .001$, two-tailed tests.

Appendix IV. Exhaustive model of the propensity score.

Proportion married and treatment effect estimates by gender and definition of marriage, using an exhaustive model of the propensity score

		ATE				ATT				ATC				HTE linear slope			
		CES-D scale		Self-rated health		CES-D scale		Self-rated health		CES-D scale		Self-rated health		CES-D scale		Self-rated health	
Definition of marriage	Prop. married	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)
Women (n = 2,602)																	
Currently married	0.70	-0.85***	(0.20)	0.15**	(0.04)	-0.75***	(0.19)	0.13**	(0.04)	-1.04***	(0.19)	0.20***	(0.04)	0.12	(0.15)	-0.01	(0.03)
Continuously married	0.52	-1.05***	(0.17)	0.15***	(0.04)	-0.87***	(0.17)	0.13**	(0.04)	-1.16***	(0.17)	0.19***	(0.04)	0.13	(0.19)	-0.01	(0.04)
Ever married	0.85	-0.46	(0.26)	0.22***	(0.06)	-0.30	(0.27)	0.20**	(0.06)	-0.74**	(0.22)	0.22***	(0.05)	0.21	(0.18)	0.05	(0.04)
Men (n = 2,684)																	
Currently married	0.69	-0.80***	(0.16)	0.17***	(0.04)	-0.69***	(0.16)	0.14**	(0.04)	-0.99***	(0.15)	0.19***	(0.04)	0.00	(0.09)	0.01	(0.03)
Continuously married	0.54	-0.65***	(0.15)	0.14***	(0.04)	-0.62***	(0.14)	0.11**	(0.04)	-0.86***	(0.15)	0.17***	(0.04)	-0.15	(0.16)	0.01	(0.05)
Ever married	0.81	-0.34	(0.19)	0.15**	(0.05)	-0.17	(0.19)	0.11	(0.05)	-0.85***	(0.18)	0.21***	(0.04)	0.32**	(0.11)	-0.01	(0.03)

* $p < .05$; ** $p < .01$; *** $p < .001$, two-tailed tests.