# Can Marriage Reduce Risky Behavior for African-Americans?

Olugbenga Ajilore[1]
Mir M Ali[2]

April 25, 2009

## Abstract

This paper estimates whether marriage can improve health outcomes for African-Americans through changes in health risky behaviors. A large part of chronic health conditions are the result of engaging in risky behaviors like smoking, drinking, and drug use. The literature has shown that marriage has positive effects on health outcomes. However, what is missing in the literature is whether this hold true for African-Americans. Our study builds upon the burgeoning literature estimating the impact of marriage on risky behavior. Using propensity score matching to account for the potential selection bias with a sample created from the National Longitudinal Study of Adolescent Health, the results show that marriage does lead to lower risky behavior, specifically drinking and drug use. We also estimate the impact of cohabitation on risky behavior and find that cohabitation is not associated with a reduction in risky behavior. This question has important policy implications because if marriage has the same benefits for African-Americans as it does for the general population, social welfare programs can be re-evaluated to incorporate marriage promotion and further support can be given to programs that decrease adverse health behaviors.

PRELIMINARY. PLEASE DO NOT CITE WITHOUT THE AUTHORS PERMISSION.

[1] Department of Economics, University of Toledo, Toledo, OH 43606-3390, USA, phone: (419) 530-2113, fax: (419) 530-7844, e-mail: gbenga.ajilore@utoledo.edu.
[2] Department of Economics, University of Toledo, Toledo, OH 43606-3390, USA, phone: (419) 530-5148, fax: (419) 530-7844, e-mail: mir.ali3@utoledo.edu

## 1. Introduction

The extensive literature on marital outcomes largely found married adults to be generally healthier than their unmarried counterparts. Married adults were documented to have lower mortality rates, lower morbidity rates and to be in better physical and mental health (Waldron et al., 1996). The absence of isolation, social support, and economic well-being has been proposed as reasons as to why marriage is positively correlated with health. (Ross, Mirowsky, and Goldsteen 1990; Coombs, 1991). Since goods and services contribute to the production of health (Grossman, 1972) and because marriage may involve transition from a one to a two person household, this may increase the resource endowments (income, health insurance) of the family thus contributing to better health. Also increased time availability in a married household, due to gains from specialization and exchange in presence of comparative advantage, would allow for greater investments in health. While health outcomes and marriage have been widely studied, the effects of marriage on unhealthy risky behavior have received relatively little attention (Duncan et al., 2006; Lin, 2008). Risky behaviors such as smoking, heavy drinking and drug use are associated with numerous detrimental health outcomes and a large part of most chronic health conditions are the result of engaging in such risky behaviors. What is missing in the literature is the effect of marriage on health outcomes among African-Americans. Waite (1995) finds marriage to have an overall positive effect on individuals, but such effects were not evenly distributed by race and gender. African-Americans have lower rates of marriage and these trends have been consistent over time (Dixon, 2009). Because of shortages of African American men relative to women, marriage market conditions are considered one possible explanation for low marriage rates among African Americans.

Imbalanced sex ratios at birth along with excess mortality rates and high incarceration rates among young African American men could have lead to such shortages in most U.S cities (Harknett and McLanahan, 2004). Lower economic stability and employment prospects are also considered to be reasons behind low marriage rates among African Americans. Beyond socioeconomic explanations, marital norms and attitudes towards marriages are also attributed for lower marriage rates among African Americans. In addition, a 2006 report released by the Center for Disease Control and Prevention (CDC) found an increasing trend of participation in health risky behaviors among African-Americans. Thus in this paper we aim to estimate whether marriage can improve health outcomes for African-Americans through changes in such risky behaviors, namely smoking, drinking, and drug use. While a risky behavior encompasses many other types of actions, we focus on these three because they have long-term impacts on health outcomes.

The literature identifies two potential mechanisms through which marriage could impact health outcomes: the marriage protection effect and the marriage selection effect (Goldman, 1993; Murray, 2000). The marriage protection effect refers to the beneficial effects that can stem from marriages, such as increased social support, increase in income, healthy lifestyle etc., all of which contribute to better health outcomes (a reduction in risky behaviors in our case). The marriage selection effect on the other hand, could occur when healthy individuals are disproportionately likely to opt into marriages. Unhealthy people are considered less desirable marriage partners (Waldron et al., 1996). In the presence of such selection effects, the estimates of marriage effects on health outcomes could be biased, since we might be measuring the decision by healthy individuals to get married rather than the better health outcomes that might result from marriages. From a policy perspective, it is important that such biases be purged from the estimates to empirically quantify the beneficial effects resulting from marriages (marriage protection effects).

We control for the selection bias by utilizing the propensity score matching technique. Propensity score matching is a two-step process whereby the conditional probability of selection into treatment (propensity scores) are created for each individual and then the treated individuals (i.e. married) get matched with the control individuals (i.e. unmarried), given a matching algorithm. From the matches, the average treatment effect on the treated is estimated. This methodology potentially mitigates the marriage selection bias. We estimate our models using data from the National Longitudinal Study of Adolescent Health (AddHealth). Being a longitudinal dataset that follows individuals from their adolescence till adulthood, it allows us to control for a wide range of variables (from both adulthood and adolescence) that potentially measures not only the propensity of marriages but also the likelihood of engagement in risky behavior. We also account for cohabitation in our models. While marriage is the most common foundation of family life in the U.S, there has been an upward trend in non-marital relationships as a process of family formation, especially cohabitation (Heiland and Liu, 2006). Although cohabitation can act as a precursor to marriage, not all cohabitations transition into marriage (Manning and Stock, 1995; Manning and Landale, 1996). Also in cohabitation, the norms for not engaging in health compromising behavior are less clear, because only in the recent decades has cohabitation become a prevalent living arrangement. This allows us to further identify whether marriage as a form of family life presents benefits that are not observed in other non-marital relationships.

Our results indicate that marriage decreases risky behavior for African-Americans, especially in the case of drug use and drinking. The results also show that cohabitation does not lead to a reduction in risky behavior. This result underscores the importance of differentiating between marriage and cohabitation, because although African-Americans have lower rates of marriage, they are often found in cohabitating relationships. The remainder of the paper is organized as follows: a detailed description of the data and the sample of our analysis are presented in Section 2; Section 3 describes our estimation strategies; Section 4 discusses our results, and finally Section 5 provides conclusions and policy implications.

## 2. Data

The data set used in our study comes from the National Longitudinal Study of Adolescent Health (AddHealth). AddHealth consists of data on U.S students in 132 schools nationwide between grades 7 to 12. AddHealth data includes three waves of in-home surveys first conducted in 1994 with follow-up surveys in 1996 and in 2002, when most respondents had made a transition to adulthood. The primary data for our analysis came from all three waves (1994, 1996 and 2002) of the in-home survey portion of AddHealth. The third wave of the data includes individuals of marriageable age. A primary advantage of the data set is its longitudinal nature, which allows us to control for past participation in risky behavior when individuals were in their adolescence. This helps us to capture the addictive nature of such risky behaviors, especially smoking. It might possibly be that individuals who were more prone to risky behavior in their adolescence continue with such participation into their adulthood. In such regards, the effect that marriage could have on risky behavior could be biased downwards. Controlling for risky behavior during adolescence allows us to obtain a more precise estimate of the marriage protection effect. The sample of our analysis includes all African - Americans that were interviewed in all three waves of study (N = 2, 581).

Using Addhealth also allows us to focus on how marriage affects behaviors among a relatively younger cohort, the life stage in which the marriage promotion policies have most recently been focused. By focusing on a younger cohort it might be possible for us to inform policy better because early experiences can substantially impact subsequent behaviors (Schoen et al., 2007). An evidence of positive health benefits from marriage among a demographic group that has a low propensity of marriage (African-Americans in early adulthood), would imply a much larger returns to health from marriage among older African Americans.

*2.1 Measures of Risky Behaviors*

The dependent variables of our analysis pertains to participation in risky behaviors, particularly, smoking, drinking and drug use. There are two smoking measures, two drinking measures and one drug use measure. The first smoking variable is an indicator variable for whether the individual smokes on a daily basis (all days out of the past 30 days) and the second variable is the number of cigarettes smoked on the day the individual smoked. The first drinking variable is a dichotomous variable indicating whether the individual has participated in heavy drinking or not. Heavy drinking is defined as having more than five drinks usually at one time for each time they had a drink in the last two weeks. The second drinking variable is also a dichotomous variable indicating whether the individual got drunk in the past 12 months. The drug use variable is a binary measure indicating whether an individual tried marijuana at least once in the last thirty days.

*2.2. Union Status Variables*

The primary variable of interest in our study is the union status of the individuals, i.e. the individual's marital or cohabitation status. Our union status variables are indicators of whether the individual is currently married or not and also whether the individual is currently cohabitating or not. It is important to analyze both cohabitation status and marriage because even though marriage and cohabitation are similar in some sense, there are behavioral differences between individuals who are cohabitating and individuals who are married (Axinn and Thornton, 1992; Rindfuss and VandenHeuvel, 1990; Winkler, 1997). For example, Rindfuss and VandenHeuvel (1990) found cohabitators behave more like singles rather than married individuals. Also Winkler (1997) found that cohabitators don't pool all their income together.

*2.3 Other Control Variables*

Besides controlling for demographic characteristics like age, gender, birth weight, nativity, and education in our models, we also control for employment status (dummies for working full-time), and current pre-tax income. Waldron et al. (1996) finds income and work-status to account for part of the marriage protection effect. Also included are measures like whether individual lived in a two-parent household, whether they have a sibling and whether they consider religion to be important to them. We also control for measures of physical and mental health during adolescence including participation in risky behaviors (from Wave II: 1996). Physical health during adolescent years are measure by controlling for obesity status measured using the CDC age and gender adjusted cutoffs in BMI (calculated using measured height and weight) and a self-reported indicator of being in good health. Our mental health indicator is based on a dichotomized version of the Center for Epidemiologic Studies Depression (CES-D) Scale, which is a very widely used measure of depressive symptom. AddHealth administered 18 of the 20 items that typically comprise the CES-D Scale. Specifically, respondents were asked to indicate the frequency with they had experienced certain feelings or emotions during the past week, like how often they felt “life had been a failure,” how often they felt “lonely,” and how often they “talked less than usual.” Possible responses were “rarely or none of the time” (=0); “some or a little of the time” (=1); “occasionally or a moderate amount of the time” (=2); and “most or all of the time” (=3). Responses to these 18 items were summed to produce a score of between 0 and 54. From this the depression indicator was set equal to 1 if a male respondent scored above 22 on the CES-D Scale and 0 otherwise. A cut-point of 24 was utilized for the female respondents (Sabia and Rees, 2008). We also include characteristics of the parent like the parents’ education, whether their parents are religious, whether the parents were on welfare and whether their parents engaged in smoking or drinking. Finally, we control for factors related to attitude towards marriage, like, whether the individual think it is important to be faithful in a marriage, whether they believe that lifelong commitment is important for marriage, whether they believe that financial solvency is important in a marriage and whether they think it is ok to live with a partner that they don’t intend to marry. Table 1 reports descriptive statistics for all variables of interest among our study sample.

## 3. Methodology

As mentioned previously, since marriage is a non-random event, a simple regression about the effects of marriage on risky behavior, may be ignoring potential selection biases. It is quite possible that individuals who engage less in risky behavior are more prone to being married. The selection issue may also have implications for African-Americans specifically, since from our descriptive statistics we see that African-Americans are less likely to be in any relationship. Therefore, it is likely that the non-random selection into marriage could complicate the estimation of the marriage effect. Simple comparisons of risky behavior outcomes by marital status can be misleading if individuals who get married are different from those who remain unmarried. A method of correcting this selection bias is to use propensity score matching (Rosenbaum and Rubin, 1983). Propensity score matching is a technique that is used to adjust for pre-treatment observable differences between a treatment group and a control group; thus the primary purpose of this methodology is "to replicate conditions of an experiment such that the treatment variable, in this case marriage, can be treated as though it occurred at random and that the individuals under analysis are homogenous on all other factors except the treatment variable" (King et al., 2007, pg. 43). Propensity score matching allows us to formulate the relationship between marriage and health risky behavior in a framework similar to a social experiment in which the treatment is randomly assigned. In our context, the treatment (marriage) is defined in terms of the potential outcomes for those who married (treated). We estimate matching methods to construct the counterfactual outcomes for the treated in the absence of a treatment, by matching the treated with controls (individuals who did not marry) who share identical characteristics that rule selection into treatment. Although this methodology addresses selection on observables, it does not extend to selection on unobservable; thus, like the literature we also rely on the richness of our data set to reduce such biases generated by unobservables.

*3.1. Empirical Framework*

Closely following the notations used by Liu and Heiland (2008) we provide an intuitive exposition of our estimation framework; consider an individual $i$ who engages in health risky behaviors $R_i$. The interrelation of the risky behavior and marital status can be presented as:

$$R_i = \beta M_i + \alpha X_i + \varepsilon_i$$
$$M_i = \eta X_i + v_i$$

where $M_i$ equals to 1 if the individual is married and 0 otherwise. Characteristics of the individual that influences their engagement in health risky activities and marital outcome is represented by $X_i$. Unobservable characteristics affecting $R_i$ and $M_i$ are captured by $\varepsilon_i$ and $v_i$. The effect of marriage on risky behavior is measured by $\beta$. However, estimating $R_i$ directly may yield a biased estimate of $\beta$ if $M_i$ and $\varepsilon_i$ are statistically dependent. Two main sources can be attributed to this dependency (Rosenbaum and Rubin, 1983; Heckman and Robb, 1985): first, $X_i$ and $\varepsilon_i$ may be correlated, (the individuals' characteristics may be correlated with unmeasured addictive propensities); second, $\varepsilon_i$ and $v_i$ may be correlated (unobserved factors may affect both risky behaviors and marital status). The existence of an either source of bias would likely show that married individuals have different outcomes compared to their non-married counterparts independent of any causal effect of marriage. Selection bias may arise in the regression analysis since these estimators would utilize data from all observations to be combined into one estimate of the marriage effect. The validity of the estimate would be suspect, if individuals who marry are different from those who don't. In the presence of any factors that affect the individuals' decision to marry as well as their engagement in risky behavior, the estimate will reflect both the marriage protection effect (the 'true' marriage effect we want to identify) and the marriage selection effect (the effect that influences the individual's decision to marry in the first place).

In our analysis the treatment is marriage, thus $M_i = 1$ denotes the treatment group and $M_i = 0$ denotes the control group (individuals who do not marry). Let $R_i(1)$ denote the potential outcome of individual $i$ under the treatment state ($M_i = 1$) and $R_i(0)$ the potential outcome if the same individual $i$ receives no treatment ($M_i = 0$). Thus $R_i = M_i R_i(1) + (1 - M_i) R_i(0)$ is the observed outcome of individual $i$. The individual treatment effect $\beta_i = R_i(1) - R_i(0)$ is unobserved since either $R_i(1)$ or $R_i(0)$ is missing. Standard parametric models (e.g. OLS) estimates the average treatment effect (ATE) by taking the average outcome difference between the treatment groups: $\beta_{OLS} = E[R_i(1) | M_i = 1] - E[R_i(0) | M_i = 0]$. If individuals who remained unmarried are unlikely to ever marry, the ATE may not be particularly helpful in understanding how marriage affects participation in risky behaviors. An alternative is to estimate the average treatment effect on the treated (ATT):

$$\beta_{M_{i=1}} = E[\beta_i \mid M_i = 1] = E[R_i(1) \mid M_i = 1] - E[R_i(0) \mid M_i = 1]$$

which is the difference between the expected outcome of an individual who marries and the expected outcome of the same individual if he/she were to remain unmarried.

While we observe the outcomes of the married individuals and thus are able to construct the first expectation $E[R_i(1) \mid M_i = 1]$, we cannot identify the counterfactual expectation $E[R_i(0) \mid M_i = 0]$ without invoking further assumptions. To overcome this problem, we have to rely on the individuals who remain unmarried to obtain information on the counterfactual outcome. A way to construct a sample counterpart for the counterfactual outcomes of the treated had they not received treatment is to use statistical matching. The matching estimators can be devised to reconstruct the condition of an experiment by stratifying the sample with respect to covariates $X_i$ that rule selection into treatment. Selection bias is eliminated provided all variables in $X_i$ are measured and balanced (comparable) between the two treatment groups within each stratum. In this case, each stratum represents a separate randomized experiment and simple outcome difference between the treated and controls provide an unbiased estimate of the treatment effect (Liu and Heiland, 2008).

An identifying assumption of the matching method is the Conditional Independence Assumption (CIA), i.e. all relevant outcome differences between the matched treated and controls are captured in their observed characteristics. Hence, conditional on $X$, the outcomes of those who remained unmarried are what the outcomes of those who married would have been if they had remained unmarried. The conditional response of the treated under no treatment could thus be estimated by the conditional mean response of the matched untreated. To estimate the ATT, one is first require taking the outcome difference between the two treatment groups conditional on $X$, then average over the distribution of the observables in the treated population. Rosenbaum and Rubin (1983) proposed using the conditional probability of selection into treatment (propensity score) to stratify the sample. They demonstrate that by definition the treated and the non-treated with the same propensity score have the same distribution of $X$ . This is also called the balancing property of the propensity score. Matching treated and untreated units using their estimated propensity score and placing them into one block (i.e. observations with propensity score falling within a specified range) means that selection into treatment within each block is random and the probability of receiving treatment within this block equals the propensity score. However, the probability of finding an exact match is theoretically zero. Thus, a certain distance between the treated and the untreated has to be accepted (Becker and Ichino, 2002). A variety of matching algorithms have been used in the literature, including Gaussian, Epanechinikov and Uniform (radius) Kernel matching, with none a priori superior than the other. Since there is no consensus in the existing literature about what the appropriate or the most efficient matching algorithm is, we utilize all of the mentioned algorithms and compare our estimates. This also provides a way to check the robustness of our results.

**4. Results**

We begin first by estimating the propensity score for selection into the treatment, by using a probit model. An important issue in implementing the probit model is to decide on the covariates to be included. Here we rely on the proposition by Rosenbaum and Rubin (1983) and Dehejia and Wahba (1999): for any given specification, group observations into blocks defined by the estimated propensity score and verify whether it succeeds in balancing the covariates between the treated and the controls within each block. If a particular structure that balances the covariates are not found (indicating that the specification does not capture the differences between the treated and the controls), we include additional covariates until this condition is satisfied. We begin by including the simplest set of controls (age, gender, education and religion) and finally succeed in balancing the covariate means when we include initial health endowments for the individuals' adolescent years. Table 2 presents results of the balancing test between the treated and the control groups after stratifying the sample into blocks based on their estimated propensity score. From the table we can see that the characteristics of the matched control within each block resemble the treated group, showing that the balancing condition is satisfied. Matching based on the full set of controls result in a sample of 2,581 observations with propensity score falling within the region of common support (0.009, 0.463). Figure 1 also shows that the treated and the control are comparable since there is sufficient overlap in the propensity score within each block. In Table 3 we present the probit estimate of the propensity score for the fully specified model.

Table 4 presents our OLS estimates along with Gaussian, Epanechinikov and Uniform (radius) Kernel matching estimates. To assess the sensitivity of the matching estimates to the choice of bandwidth (or radius) we report results using different bandwidths. We report estimates for our main variable of interest only. However, it is important to note that our model controls for past participation in risky behaviors besides other demographic, health and parental characteristics. An important result from our estimation is that participation in risky behavior at the baseline is positively related to current risky behaviors (not reported) and these effects are the largest in magnitude among all the covariates.  This highlights the needs for controlling past risky behavior since for most individuals engagement in risky behavior could be habitual. Not controlling for this could have provided us with an overestimation of the effect of marriage.

Our OLS estimates show that African-Americans who are married are less likely to engage in heavy drinking, getting drunk and drug use. On average individuals who are married are 4.4% less likely to engage in heavy drinking, 10% less likely to get drunk and 5.3% less likely to use any drugs. While the matching estimates confirm the direction of the effect implied by the parametric results, they suggest that being married reduces the individual's likelihood of engaging in risky behavior relative to if the individuals remained unmarried (marriage protection effect). The coefficients are also larger in magnitude compared to the OLS estimates; however, the effect of marriage on heavy drinking is only significant in two of the matching algorithms. The finding that, on average, the outcome difference between a given treated individual and an individual in the control group is smaller (OLS) than the outcome difference between the same treated individual and an individual in the control group (PSM) who exhibit similar characteristics implies that marriage yields some protective benefits. Similar to the models estimated without propensity scores, marriage did not have a statistically significant effect on either the probability of being a regular smoker or on the number of cigarettes smoked. The insignificant effect of marriage on smoking could possibly be due to the fact that smoking may not be considered necessarily as a detrimental behavioral outcome that requires immediate curtailing upon marriage as opposed to heavy drinking, getting drunk or drug use. Also the negative health outcomes of smoking tend to unfold more in the long run in addition to it being related to a high level of dependency or addiction. The effects of marriage on smoking maybe a longer run phenomenon and may require a longer panel of data to be estimated empirically.

As discussed previously, marriage may induce better health outcomes (reduction in risky behavior) for a number of reasons including gains from specialization and social support. Differences in health endowments or background characteristics are less plausible explanations since we match married individuals with unmarried individuals who are similar in these characteristics. Given that some of the benefits to marriage could extend to cohabitating individuals as well (e.g. economies of scale), we re-estimate the models with cohabitation as the control group. These results reported in Table 5 indicate better outcomes for the treated (married) in terms of heavy drinking, getting drunk and drug use. This pattern is consistent with the idea that in cohabitation the norms of non-engagement in health compromising behaviors are less clear (due to greater instability) and also individuals in a cohabitating relationship behaves more like singles (Rindfuss and VandenHeuvel, 1990). Thus gains from non-marital unions are less compared to marital unions.

## 5. Conclusion

Our results indicate that marriage among African Americans may lead to a reduction in risky behaviors, especially for alcohol consumption and drug use. Such benefits are also greater compared to those who are in a cohabiting relationship. Our models are analyzed after accounting for selection bias, thus it can be suggested that marriage may exhibit a protection effect, but such effects are not present in cohabitation.

Marriage is governed by social norms which are prevalent enough to be called an ‘institution’ and thus could possibly entail to ‘cleaning up one’s act’ (Duncan et al., 2006). Cohabitation on the other hand may be viewed as an incomplete institution, where the norms are less clear. So even if cohabitation can lead to some reduction in risky behavior, the norms are less agreed upon (Hofferth, 2006). This could potentially offset the 'monitoring of partners' potential of the co-residence, since cohabitation unlike marriage does not involve more engagement with one's partner and are also not a long-term commitment like marriage.

The insignificant effect of marriage on smoking is similar to the previous literature (Duncan et al., 2006), and could be a result of the high level of dependence of smokers on cigarettes (addiction) or may be because smoking is not yet considered as something that one has to give up after marriage like heavy drinking and drug use. It could also be a result of smokers selecting other smokers as partners (Clark and Elite, 2006) to a higher degree than alcohol or drug use. This brings us to a potential short coming of the paper, which is the possibility of homophily or positive assortative mating on risky behaviors. Individuals who smoke or like to drink or use drugs might be more likely to marry or live with a person with similar preferences. However, since AddHealth does not contain information on spouse's behavior, we are unable to control for this.

From a policy perspective, our results indicate the positive or the beneficial effect that marriage can have among African Americans in terms of reducing participation in risky behaviors. This positive benefit was not evident in other forms of union, like cohabitation. A possible extension of this study could be to analyze the effect of marriage on diet and exercise behaviors. This is especially relevant among African Americans given a higher percentage of diabetes, heart disease and obesity among them.

**References**

Axinn, William G. and Thornton, A. 1992. “The Relationship Between Cohabitation and Divorce: Selectivity or Causal Influence?” *Demography* 29(3): 357-374.

Becker, Sascha & Andrea Ichino, (2002). "Estimation of average treatment effects based on propensity scores," *Stata Journal*, StataCorp LP, vol. 2(4), pages 358-377.

Caliendo, Marco and Sabine Kopeining (2008). “Some Practical Guidance for the Implementation of Propensity Score Matching,” *Journal of Economic Surveys* 22(1): 31–72

Clark, Andrew and Fabrice Etile (2006) “Don’t Give Up On Me Baby: Spousal Correlation in Smoking Behaviour” *Journal of Health Economics* 25: 958–978

Coombs, Robert (1991) “Marital Status and Personal Wellbeing: A Literature Review” *Family Relations* 40: 97–102

Dehejia, R.H., and S. Wahba (1999) “Causal Effect in Nonexperimental Studies: Reevaluation of the Evaluation of Training Programs.” *Journal of the American Statistical Association*, 94(488), 1053–62.

Duncan, Greg, Bessie Wilkerson, and Paula England (2006) “Cleaning Up Their Act: the Effects of Marriage and Cohabitation on Licit and Illicit Drug Use” *Demography* 43(4): 691–710

Goldman, N. (1993) “Marriage Selection and Mortality Patterns: Inferences and Fallacies”, *Demography* 30(2): 189 – 208.

Grossman, M. (1972) “On the Concept of Health Capital and the Demand for Health”, *Journal of Political Economy* 82: 223 – 255.

Harknett, K., McLanahan, S. (2004) “Racial and Ethnic Differences in Marriage after the Birth of a Child”, *American Sociological Review* 69(6): 790 – 811.

Hayward, Mark, Toni Miles, Eileen Crimmins, and Yu Yang (2000) “The Significance of Socioeconomics Status in Explaining the Racial Gap in Chronic Health Conditions” *American Sociological Review* 65(6): 910–930

Heiland, F and Liu, S. (2006) “Family Structure and Wellbeing of Out-of-Wedlock Children: The Significance of the Biological Parents’ Relationship”, *Demographic Research* 15: 61 – 104.

Heckman, J.J., and R. Robb (1985) “Alternative Methods for Evaluating the Impact of Interventions: An Overview.” *Journal of Econometrics*, 30(1), 239–67.

Hofferth, S. (2006) “Residential Father Family Type and Child Well-being: Investment Versus Selection.”*Demography* 43:53–77.

King, Ryan, Michael Massoglia, and Ross MacMillan. (2007) “The Context of Marriage and Crime: Gender, the Propensity to Marry, and Offending in Early Adulthood” *Criminology* 45: 33–65

Kington, Raynard and Herbert Nickens (2001) “Racial and Ethnic Differences in Health: Recent Trends, Current Patterns, Future Directions” in eds. Neil Smelser, William Junius Wilson, Faith Mitchell, *American Becoming: Racial Trends and Their Consequences* (National Research Council) pg. 253–310

Leuven, Edwin and Barbara Sianesi (2003) “PSMATCH2: State Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance testing” http://ideas.repec.org/c/boc/bocode/s432001.html/

Lin, Muh-Ching (2008) “Unhealthy Behavior and Marriage” *Working Paper*

Liu, S and Heiland, F (2008) "Should We get Married? The Effect of Parent's Marriage on Out-of-Wedlock Children" *Economic Inquiry*, Forthcoming.

Manning, Wendy and Nancy Landale (1996) “Racial and Ethnic Differences in the Role of Cohabitation in Premarital Childbearing” *Journal of Marriage and the Family*, 58(1): 63–77

Manning, Wendy and Pamela Stock (1995) “Why Marry? Race and the Transition to Marriage among Cohabitors” *Demography* 32(4): 509–520

Murray, John E. (2000) “Marital Protection and Marital Selection: Evidence from a Historical-Prospective Sample of American Men.” *Demography*, 37(4): 511–521

Raley, R. Kelly and Megan Sweeney (2007) “What Explains Race and Ethnic Variation in Cohabitation, Marriage, Divorce, and Nonmarital Fertility?” *California Center for Population Research Working Paper* CCPR-026-7

Reynolds, C. Lockwood and Stephen L. DesJardins. 2008. “The Use of Matching Methods in Higher Education Research: Answering Whether Attendance at a Two-Year Institution Results in Differences in Educational Attainment” in Higher Education: Handbook of Theory and Research, Volume XXIII, J.C. Smart (Ed.), *Forthcoming*.

Rindfuss, Ronald R. and Audrey VandenHeuvel. 1990. “Cohabitation: A Precursor to Marriage or an Alternative to Being Single?” *Population and Development Review* 16: 703-726.

Rosenbaum, Paul R. and Donald B. Rubin (1983) “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika* 70(1): 41–55

Ross Catherine, John Mirowsky, and Karen Goldsteen (1990) “The Impact of the Family on Health: The Decade in Review” *Journal of Marriage and the Family* 52: 1059–1078

Sabia, Joseph J. and Daniel I. Rees (2008) "The Effect of Adolescent Virginity Status on Psychological Wellbeing." *Journal of Health Economics* 27(5):1368-1381.

Schoen R, Landale NS and Daniels, K (2007) “Family Transitions in Young Adulthood” *Demography* 44(4): 807 – 820.

Schoenbaum, Michael and Timothy Waidmann (1997) “Race, Socioeconomic Status, and Health: Accounting for Race Differences in Health” *The Journal of Gerontology* Series B 52B: 61–73

Waite, Linda (1995) “Does Marriage Matter? “ *Demography* 32(4): 483–507

Waldron, I., Hughes, M.E., and Brooks, T.L (1996) “Marriage Protection and Marriage Selection – Perspective Evidence for Reciprocal Effects of Marital Status and Health” *Social Science and Medicine*, 43(1): 113 – 123.

Wilson, Chris M. and Oswald, Andrew J., (2005) "How Does Marriage Affect Physical and Psychological Health? A Survey of the Longitudinal Evidence" *IZA Discussion Paper No. 1619*

Winkler, Anne E. 1997. “Economic Decision-Making by Cohabitors: Findings Regarding Income-Pooling.” *Applied Economics* 29(8): 1079-1090.

**Table 1: Descriptive Statistics**

| Variable (N = 2,581) | Mean | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|
| **Outcome** | | | | |
| Daily Smoker | 0.098 | 0.297 | 0 | 1 |
| Number of Cigarette Smoked | 3.972 | 5.129 | 0 | 100 |
| Heavy Drinker | 0.184 | 0.387 | 0 | 1 |
| Drunk | 0.285 | 0.951 | 0 | 1 |
| Drug | 0.191 | 0.393 | 0 | 1 |
| **Union Status** | | | | |
| Married | 0.087 | 0.282 | 0 | 1 |
| Cohabit | 0.129 | 0.335 | 0 | 1 |
| **Health Measures (Wave II: 1996)** | | | | |
| Birth Weight | 6.344 | 1.262 | 3 | 12 |
| Obese | 0.155 | 0.362 | 0 | 1 |
| Self- Reported Good Health | 0.928 | 0.258 | 0 | 1 |
| Depressed | 0.084 | 0.278 | 0 | 1 |
| Smoke | 0.886 | 0.317 | 0 | 1 |
| Drink | 0.336 | 0.572 | 0 | 1 |
| Drug | 0.135 | 0.342 | 0 | 1 |
| **Demographics** | | | | |
| Male | 0.433 | 0.496 | 0 | 1 |
| Age | 21.56 | 1.656 | 18 | 27 |
| College | 0.077 | 0.267 | 0 | 1 |
| Has Siblings | 0.384 | 0.486 | 0 | 1 |
| Born USA | 0.767 | 0.423 | 0 | 1 |
| Log Income | 8.380 | 2.181 | 0 | 12.924 |
| Religious | 0.859 | 0.348 | 0 | 1 |
| **Parental Characteristics (Wave I: 1994)** | | | | |
| Lived with both Biological Parents | 0.322 | 0.467 | 0 | 1 |
| Parent Black | 0.802 | 0.398 | 0 | 1 |
| Parent College | 0.129 | 0.336 | 0 | 1 |
| Parent Religious | 0.129 | 0.336 | 0 | 1 |
| Parent Smoke | 0.239 | 0.427 | 0 | 1 |
| Parent Drink | 0.522 | 0.499 | 0 | 1 |
| Welfare | 0.359 | 0.480 | 0 | 1 |
| **Attitude Towards Marriage** | | | | |
| Faithfull | 0.862 | 0.345 | 0 | 1 |
| Life Long | 0.651 | 0.476 | 0 | 1 |
| Finance | 0.369 | 0.483 | 0 | 1 |
| Live Together | 0.651 | 0.477 | 0 | 1 |
| | | | | |

**Table 2: Test of Balancing Properties Between the Control and Treated Group (Two-Sample T Test of Means): T-Statistic Reported**

| | **Block 1** | **Block 2** | **Block 3** | **Block 4** | **Block 5** |
|---|---|---|---|---|---|
| N Treated | 18 | 59 | 96 | 48 | 4 |
| N Control | 839 | 694 | 530 | 163 | 4 |
| Range of the Propensity Score | [0.009, 0.050] | [0.050, 0.010] | [0.100, 0.200] | [0.200, 0.400] | [0.400, 0.463] |
| | Two-Sample Test of Means: \|T\| Statistics | | | | |
| Propensity Score | 0.1217 | 0.3337 | 0.1304 | 2.0830 | 0.2888 |
| Male | 0.6515 | 1.0761 | 0.9038 | 1.5924 | 0.6547 |
| Age | 0.3282 | 0.2965 | 0.7870 | 1.3983 | 0.8783 |
| College | 1.2585 | 0.6503 | 1.4739 | 0.5417 | 0.6000 |
| Has Siblings | 1.6678 | 0.4569 | 0.9700 | 1.0080 | 1.7321 |
| Born USA | 0.7641 | 0.7057 | 2.2056 | 2.1837 | 0.6077 |
| Log Income | 0.9564 | 0.0366 | 0.1881 | 0.6583 | 0.2015 |
| Religious | 1.8965 | 0.4307 | 1.2103 | 0.6583 | 0.6006 |
| Obese | 0.3905 | 1.6009 | 0.8649 | 1.2711 | 0.6547 |
| Birth Weight | 1.3188 | 0.3149 | 1.0424 | 0.9969 | 0.0493 |
| Self- Reported Good Health | 0.0716 | 0.7378 | 1.2458 | 1.6867 | 1.7321 |
| Depressed | 0.9969 | 1.2485 | 0.6636 | 0.8020 | 0.6547 |
| Smoke | 1.4535 | 0.9419 | 0.5613 | 0.7941 | 1.0006 |
| Drink | 1.0196 | 0.9764 | 0.4064 | 1.4555 | 0.2064 |
| Drug | 0.3170 | 0.7596 | 0.0340 | 0.8548 | 1.0006 |
| Lived with both Biological Parents | 0.1086 | 0.3602 | 1.3488 | 1.7236 | 0.6007 |
| Parent Black | 0.1705 | 0.7058 | 0.7489 | 0.5830 | 1.0086 |
| Parent College | 1.7228 | 2.2937 | 1.6666 | 2.5438 | 0.0514 |
| Parent Religious | 0.8419 | 0.1203 | 0.1384 | 1.4171 | 1.0054 |
| Parent Smoke | 0.3997 | 1.4740 | 1.0769 | 0.9778 | 0.0947 |
| Parent Drink | 0.0714 | 0.5818 | 0.6468 | 0.3531 | 0.3871 |
| Welfare | 0.2952 | 0.0448 | 0.4008 | 1.1980 | 0.0006 |
| Faithfull | 0.5262 | 0.0809 | 0.8303 | 1.6178 | 0.0600 |
| Life Long | 1.1343 | 0.5889 | 0.5481 | 0.5417 | 0.5017 |
| Finance | 0.0992 | 1.7737 | 0.7168 | 2.0731 | 1.7321 |
| Live Together | 0.2014 | 0.3418 | 1.5616 | 2.1004 | 0.6547 |
| | | | | | |

**Table 3: Probit Estimates of the Propensity Score**

| Variables | Coefficient | Standard Error | P > \|z\| |
|---|---|---|---|
| | | | |
| Male | -0.112 | 0.079 | 0.159 |
| Age | 0.221 | 0.025 | 0.000 |
| College | -0.431 | 0.153 | 0.005 |
| Has Siblings | 0.084 | 0.079 | 0.289 |
| Born USA | 0.026 | 0.092 | 0.775 |
| Log Income | 0.019 | 0.017 | 0.282 |
| Religious | 0.099 | 0.116 | 0.395 |
| Obese | -0.079 | 0.108 | 0.467 |
| Birth Weight | 0.073 | 0.031 | 0.018 |
| Self- Reported Good Health | -0.050 | 0.141 | 0.723 |
| Depressed | 0.152 | 0.124 | 0.221 |
| Smoke | -0.019 | 0.122 | 0.877 |
| Drink | 0.009 | 0.007 | 0.229 |
| Drug | -0.088 | 0.113 | 0.438 |
| Lived with both Biological Parents | -0.031 | 0.087 | 0.718 |
| Parent Black | -0.127 | 0.132 | 0.336 |
| Parent College | 0.010 | 0.117 | 0.930 |
| Parent Religious | 0.133 | 0.148 | 0.369 |
| Parent Smoke | 0.050 | 0.094 | 0.596 |
| Parent Drink | -0.039 | 0.085 | 0.641 |
| Welfare | -0.061 | 0.087 | 0.480 |
| Faithfull | 0.191 | 0.151 | 0.207 |
| Life Long | 0.531 | 0.126 | 0.000 |
| Finance | -0.184 | 0.080 | 0.022 |
| Live Together | -0.013 | 0.079 | 0.873 |
| | | | |
| | Log Likelihood = -686 | | |
| | Pseudo R2 = 0.101 | | |
| | N = 2,581 (Treated = 225, Control = 2,356) | | |
| | | | |

**Figure 1: Box Plot of the Estimated Propensity Score for the Treated Units (1) and the Control Units (0) within the Common Support Region**

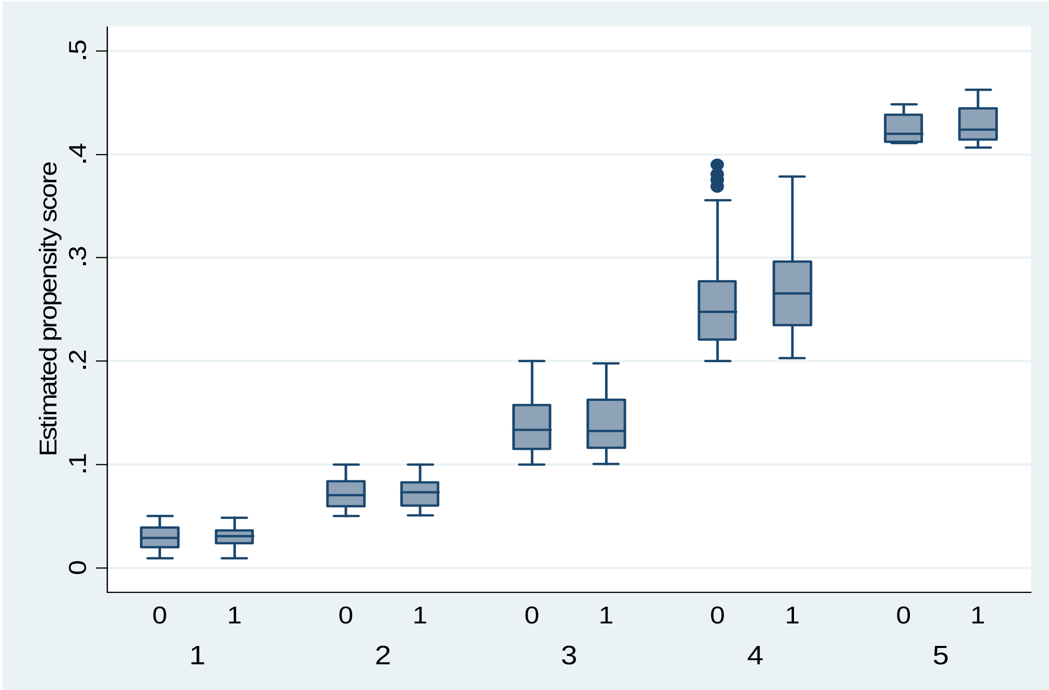


**Table 4: Estimated Effect of Marriage on Health Risky Behaviors (ATT)**

| | Parametric Estimate | Matching | | | | |
|---|---|---|---|---|---|---|
| | | | Epanechnikov | | Uniform | |
| | **OLS** | **Gaussian** | **h = 0.01** | **h = 0.005** | **r = 0.01** | **r = 0.005** |
| Daily Smoker | -0.011<br>(0.020) | -0.011<br>(0.021) | -0.008<br>(0.022) | -0.004<br>(0.023) | -0.007<br>(0.019) | -0.005<br>(0.019) |
| Number of Cigarettes | 0.439<br>(0.478) | 0.360<br>(0.419) | 0.418<br>(0.437) | 0.432<br>(0.502) | 0.561<br>(0.436) | 0.567<br>(0.431) |
| Heavy Drinker | -0.044*<br>(0.026) | -0.043*<br>(0.026) | -0.051*<br>(0.029) | -0.045<br>(0.030) | -0.037<br>(0.026) | -0.038<br>(0.025) |
| Drunk | -0.100***<br>(0.028) | -0.099***<br>(0.027) | -0.106***<br>(0.030) | -0.097***<br>(0.033) | -0.106***<br>(0.031) | -0.104***<br>(0.031) |
| Drug | -0.053**<br>(0.024) | -0.059**<br>(0.023) | -0.053**<br>(0.026) | -0.054**<br>(0.028) | -0.060**<br>(0.025) | -0.062**<br>(0.026) |
| | | | | | | |
| N Treated | 225 | 225 | 225 | 225 | 225 | 225 |
| N Controls | 2,356 | 2,356 | 2,356 | 2,356 | 2,356 | 2,356 |

*Notes*: [a] Control Group: All Unmarried; [b] Robust standard errors (OLS) and bootstrapped standard errors (Matching) are reported in parenthesis. Standard errors for the matching estimators are obtained by bootstrap with 500 replications; [c] Statistical significance at the *** = 1% level, ** = 5% level, * = 10% level. [d] Sets of controls include all variables from Table 1.

**Table 5: Estimated Effect of Marriage on Health Risky Behaviors (ATT)**

| | Matching | | | | |
|---|---|---|---|---|---|
| | | Epanechnikov | | Uniform | |
| | **Gaussian** | **h = 0.01** | **h = 0.005** | **r = 0.01** | **r = 0.005** |
| Daily Smoker | -0.040<br>(0.029) | -0.035<br>(0.034) | -0.026<br>(0.038) | -0.063<br>(0.033) | -0.055<br>(0.037) |
| Number of Cigarettes | 0.531<br>(0.476) | 0.569<br>(0.524) | -0.783<br>(0.649) | 0.433<br>(0.566) | 0.586<br>(0.612) |
| Heavy Drinker | -0.097**<br>(0.038) | -0.094**<br>(0.046) | -0.085*<br>(0.049) | -0.098**<br>(0.042) | -0.104**<br>(0.043) |
| Drunk | -0.083**<br>(0.039) | -0.081*<br>(0.046) | -0.081*<br>(0.057) | -0.075**<br>(0.044) | -0.095**<br>(0.050) |
| Drug | -0.114***<br>(0.039) | -0.109**<br>(0.044) | -0.085*<br>(0.057) | -0.104***<br>(0.040) | -0.100**<br>(0.047) |
| | | | | | |
| N Treated | 225 | 225 | 225 | 225 | 225 |
| N Controls | 333 | 333 | 333 | 333 | 333 |

*Notes*: [a] Control Group: All Cohabitators; [b] Bootstrapped standard errors (Matching) are reported in parenthesis. Standard errors for the matching estimators are obtained by bootstrap with 500 replications; [c] Statistical significance at the *** = 1% level, ** = 5% level, * = 10% level. [d] Sets of controls include all variables from Table 1.