UNHEALTHY BEHAVIOR AND MARRIAGE

MUH-CHUNG LIN

DEPARTMENT OF SOCIOLOGY

THE UNIVERSITY OF CHICAGO

UNHEALTHY BEHAVIOR AND MARRIAGE

Abstract

The gains on health for the married people have long been documented in the social sciences. Nevertheless, the precise mechanisms through which marriage improves health are rarely explored. This study examines how marriage influences health by shunning unhealthy behaviors—reductions in excessive drinking. Using data from the National Longitudinal Survey of Youth 1997 (NLSY97), we are able to control for a large number of baseline variables. Additionally, we calculated the propensity scores and used them in the frameworks of matching. We performed diagnostics to assess the validity of propensity scores and the quality of matching. After procedures to account for selection, results suggest that married people are less likely to engage in heavy drinking. The types of marriage matter: formal marriage has stronger effects, while cohabitation is unrelated to any reduction in heavy drinking.

1. Introduction

Marriage is wholesome. That married people have lower morbidity/mortality has already been observed for centuries. Married people also enjoy better mental health, have a fatter wallet, maintain a larger social network, and score higher on a wide variety of indicators of general well-being than their unmarried counterparts.

All these have been extensively documented and analyzed throughout social sciences. Lesser-known are the impacts marriage has upon unhealthy, risky, and anti-social behaviors. The reduction in these activities should be of great interests in and of itself, because such behaviors pose serious threats to the health of the population. Moreover, although sociologists and epidemiologists have long studied the mortality benefits of marriage, the exact mechanisms still elude scholars. The curtailment of unhealthy behaviors upon marriage thus could be regarded as a channel through which marriage enhances physical health and reduces morbidity and mortality.

This study intends (1) to fill in the gap in the studies of the mechanisms that generate health in household, and (2) to examine the changes in unhealthy behaviors such as hazardous drinking following changes in marital status.

2. Theoretical Perspectives

2.1 Social control and specialization

Health is a major commodity produced within the family. Families produce not only tangible goods such as meals and clothing, they also generate intangible commodities such as intimacy, comfort, and physical and mental health. In terms of household production function, when one's marital status changes, one's health also changes because of major changes in its functional form and available inputs of a household production function of health.

The functional form of household health production changes for two reasons: (1) social integration and (2) technology change. Becoming married indicates one's willingness to assume new responsibilities to oneself, to one's spouse or partner, and to the maintenance of the family. One finds new meanings in life or one's life has become more meaningful than in the state of being single. In other words, one is better integrated into the society—a Durkheimian approach articulated in Berkman et al. (2000) and Umberson (1987, 1992). Further, as Akerlof (1998) pointed out, marriage signals a rite of passage in the course of life and it "redirects the energies of the bride and the groom suggested by this [Christian sacramental] ideal. In this view, men settle down upon marriage, and if men fail to marry, they fail to settle down." Whatever channels through which social integration operates, a newly-wed person will take greater care of his/her own health thereafter and incorporate his/her spouse's expectations as new guidelines of behaviors. As a result, s/he would abstain from or cut back on unhealthy behaviors such as drinking.

Also, the functional form of household health production changes because of technology change. Prior to marriage, each spouse spreads his/her time thinly over many activities. According to Becker (1991), marriage opens the door to greater specialization because there is simply one more person. One can thus concentrate on fewer activities and trade with his or her spouse/partner. As illustrated in Rosen (1983), indivisibility implies fixed-costs elements in production and hence "rate of return is increasing in the utilization of human capital and maximized by using specialized skills as intensively as possible." As a result, for each spouse, productivity per unit of time and therefore its price increases upon becoming married. Note that we do not have to resort to the marriage wage premium for a higher price of time after marriage—for the wage premium comes after household production function changes. The price of time also increases from being single to being married because time with a spouse is more treasurable than time alone in front of a TV set. Whatever its sources, a higher price of time implies activities that consume more time will be discouraged upon marriage.

Specialization also helps produce health. In the model of traditional gendered division of labor, men and women specialize in different spheres of production: men in the labor market and women in the household. This model could be easily extended to incorporate household production of health: men use earnings from the labor market to trade with care (and nagging) from wives, while women use financial resources provided by their husbands to purchase better health—this is exactly what Lillard and Waite (1995) found.

Household production of health also changes due to changes in available inputs. Prior to marriage, there is only one person in the function. Afterwards, there is one more person to care for your own health and to exercise social control. But we must ask exactly why should one fuss over the health of his/her spouse/partner? The reason is simple: within a family, health is a public good. There are no two healths within a household—no *his* or *her* health—there is only one health. It is so not only because spouses share everything together, from the air breathed, the food eaten, to the living space occupied and the neighborhood and community (Wilson 2002), but also because the well-being of spouses is interdependent. Should one's husband/wife fall ill, s/he must sacrifice her/his own recreational or even productive activities to care for a spouse/partner. In other words, a spouse's illness compromises the well-being of two, not one, and everyone has incentives to monitor and sanction spouse's health behaviors. These social controls include promoting healthy behaviors such as maintaining an orderly lifestyle—regular exercises, eating fruit and vegetable—and discouraging health-compromising behaviors such as substance use and irregular schedules.

Moreover, one benefits from spouse's/partner's better health not only because of altruism and the aforementioned interdependency, good health is the bedrock on which human capital can accumulate, thereby enhancing productivity within family and in the labor market. One major argument against substance use is it interferes with works, studies, and functioning in daily life. Consuming alcohol or narcotics makes one drunk or stoned and unable to concentrate for hours, and the hangover and residual high also persist for a long time. In other words, those who drink or do drugs will have their flows of time disrupted frequently, which is deleterious to human capital accumulation because the fixed-cost components (getting prepared and focused) in its utilization is increased substantially.

There are many good reasons why spouses are better equipped than anyone else—including one's own parents—to promote healthy behaviors, discourage bad habits, and regulate diet and routines. In addition to the incentives discussed earlier, spouses live no more than a few feet from each other, which facilitates monitoring, and it is extremely difficult, if not impossible, to hide consumption of alcohol or narcotics from one's spouse. A smell or a look can easily reveal what one has done. Moreover, sanctions from a loved one are both most embarrassing and very difficult to withstand.

Thus far we have discussed why people reduce substance use upon entry into a marriage (or in the opposite direction once a marriage dissolves). We must ask: Is this a change in preferences caused by marriage itself? Partly yes – marriage is a major event in life course, and events as such can either reveal aspects of life previously unseen or neglected or prompt people to probe themselves deeper. As a result, preferences could change. On the other hand, the answer is partly no. First, people may well look forward to cleaning-up their own behaviors and see marriage as a window of opportunities for change. In addition, as previously discussed, to the extent that both the functional form and the inputs available for the household production function of health change following a transition into marriage, and the price of time is higher after marriage, the behaviors observed could be regarded as natural results of them.

2.2 A note on cohabitation

In a limited sense, a cohabitation union resembles a marriage. A cohabiting partner has similar accesses to monitor another's behaviors, and oftentimes has the incentives to do so. Indeed, some researchers (e.g. Willis and Michael 1994) argued that for recent cohorts, cohabitation served as a trial marriage and those who lived were no different from others. Seen in this light, we should find that those currently cohabiting people should be closer to the married in terms of unhealthy behaviors such as drinking and smoking than those who are not. On the other hand, the bulk of evidences thus far appear to suggest that cohabitors behave more like single people than the married, as suggested by Rindfuss and VandenHeuvel (1990). For instance, Winkler (1997) found that cohabitors do not pool all their income together. Further, cohabitation experiences seem to socialize one with the transience, rather than the permanence, of an intimate union (Axinn & Thornton 1992), and married people with prior cohabitation experiences are more likely to divorce, although the tendency has abated somewhat for more recent cohorts (Teachman 2003) and Lillard, Brien, and Waite (1995) demonstrated the importance of selection on this issue. Consequently, we anticipate that as a marital status group, cohabitors should have a prevalence of alcohol use that is higher than the married but lower than the currently non-cohabiting people.

2.3 Selection

These salutary effects of marriage have attracted a fair amount of criticisms on grounds of selection. Critiques argue that it is not marriage that makes people healthy and wealthy but rather it is the healthy and the productive that select into marriage. For instance, Goldman and her colleagues (1990, 1993) found that never-married singles in pre-war Japan suffered extraordinarily high mortality and, given the near universality of marriage, the unmarried indeed had astonishingly high rates of mental and physical illnesses. Goldman herself (1993, 2001) argues for a strong presence of selectivity but also pointed out that selection may take many forms such that its detection becomes very difficult. On the other hand, Lillard and Waite (1995) demonstrated the primary existence of causality despite a minor role of selection.

In our study, selection may operate in several ways. First, the marital status gradient in excessive alcohol consumption may stem from a higher propensity of the well-behaved to marry instead. People free from alcohol, tobacco, and other substances, however dull, are likely to make better spouses/companions in the long-term because they are more responsible people. If this is true, then the benefits of marriage on reduced substance use should come largely from more sober people select into matrimony and, once we control for the selection factor, there should be no marital status gap in smoking and drinking.

Selection may operate in another direction. Just as Lillard and Panis (1996) showed that people with poor health select into marriage to enjoy the related health benefits, we may see people with high risks of smoking, drinking, and substance use marry not because they want to take advantage of it but because they are immature and impulsive. Nevertheless, despite its multifarious benefits, marriage remains a risky enterprise. Should one enter it unprepared, more often than not he or she hurts both him/herself and the other dear to him/her. People who married early are often characterized as lacking maturity and acting impulsively, and these are exactly the same traits of those who score low on non-cognitive skills and who engage frequently in unhealthy behaviors studied here. If negative selection is present, we should observe that entry into marriage is positively related to hazardous drinking, or at least its salutary effects should be attenuated somewhat.

Nevertheless, due to the complexity of the issue of selection, its correction will be limited in scope. We will first employ as many baseline controls as possible to see if the effects of marriage on hazardous drinking remain. We then calculate propensity scores of heavy drinking and use them in a framework of matching to see how much of the benefits of marriage remain. In addition, we perform several diagnostics to test the validity of the propensity scores. Admittedly, these practices are limited in scope and we cannot claim to have completely purged selectivity bias from our estimates. We shall leave extensive selectivity correction for the next draft.

2.4 Homophily or positive assortative mating on drinking

It is important that we recognize the possibility of homophily or positive assortative mating on drinking. A heavy drinker is more likely to marry or live with another person with similar tastes for alcohol, if only to avoid daily clashes over the issue of drinking. Drinking together with a spouse/partner sharing the same appetite clearly make drinking even more enjoyable. In that case, marriage is not necessarily associated with a reduction in drinking and may even increase it. However, NLSY97 contains no information on spouse's drinking behaviors, and we therefore cannot pursue this line of inquiry.

3. Literature Review

The relationship between marriage and unhealthy behaviors such as hazardous alcohol consumption is not as extensively investigated as those between marriage and mortality and morbidity. Compared to studies on marriage and physical health or mental health or psychological well-being, both of which have dozens of papers, there are only about a dozen directly on unhealthy behaviors.

Umberson (1987, 1992) discussed at length how social control, operating via external and self-regulation on health behaviors, influences one's health within marriage. She found that the married were indeed more risk-averse, less likely to smoke, drink, or use substances, and more likely to maintain an orderly lifestyle. Yet, her empirical works are based on cross-sectional descriptive regressions and have not corrected for selection. Berkman, Glass, Brissette, and Seeman (2000) focused on how social integration via social networks helped to depress unhealthy behaviors.

Franks, Pienta, and Wray (2002) found smoking cessation worked only if both spouses quitted, and formerly married people smoked more. This illuminates the importance of homophily or positive assortative mating in (quitting) smoking. Power, Rodgers, and Hope (1999) found that (1) married men, and to a lesser extent married women, exhibited a larger decline in excessive drinking compared to their never-married counterparts. (2) The never-married men and women maintained a high level of alcohol consumption over time. (3) The effect of differential selection into marriage is minimal: heavy-drinking men were more likely to marry than others. (4) Divorced men and women exhibited more heavy drinking, especially in the immediate aftermath of divorce.

Martino, Collins, and Ellickson (2004) found the relationship between substance use and early marriage to be all accounted for by demographics and other mediators such as fewer years of schooling. Using NLS Labor Market Experience, Kaestner (1997) found that drug use is related to delay into first marriage and to shorter durations of marriage (higher propensity to divorce) for non-black males. For black males, substance use does not appear to influence marital choices.

Akerlof (1998), from an anthropological-economic perspective, treated marriage as a rite of passage in the life course and found married men settled down and were less likely to engage in various unhealthy behaviors, showed lower criminality, and exposed themselves in dangerous conditions less frequently.

If marriage inhibits unhealthy behaviors and risk-taking, it should also apply to couples of the same sex by reducing substance use, which more often than not precedes unprotected sex, and most importantly by curtailing the number of sex partners. After all, a marriage is a marriage. Dee (2006), using different timing of the passage of same-sex marriage laws in different countries in Europe as a natural experiment, found that the legalization of same-sex marriage did significantly reduce syphilis rates, but much less so for gonorrhea and HIV.

Duncan, Wilkerson, and England (2006) is by far substantively the most comprehensive and technically the most sophisticated study on the topic of marriage and unhealthy and risky behaviors. It is also the one closest to the current study. Using NLSY79, they investigated the changes in unhealthy and risky behaviors in a 12-month window prior and subsequent to marital status change. They found that (1) men reduced binge drinking and marijuana use after marriage, but not after entering into a cohabiting union. (2) Women decreased their heavy drinking and smoking on entry into marriage and cohabitation, but their marijuana use did not. (3) Neither men nor women changed their smoking behaviors upon entry into marriage or a cohabiting union. (4) These behavioral changes are more pronounced the closer one was about to marry or cohabit, suggesting some sort of "cleaning up one's own acts". These findings are consistent with our conjectures that unhealthy habits that interfere with workplace performance such as binge drinking and marijuana use will see a greater reduction upon marrying, while those that do not, such as smoking, are better tolerated by spouses. Further, their finding that cohabitation is more weakly related to these behavioral changes is also consistent with our conjecture based on a lesser extent of social control exerted by a cohabiting partner.

We should note that most of these papers, including Duncan et al., did not explicitly account for the issue of selection, and this is exactly where we hope this study to contribute to the literature.

4. Data and Measurement

Data: We use NLSY97 to study the relationship between hazardous drinking and marriage for its richness in background information. NLSY97 is a dataset that collects extensive information on 8,984 adolescents born between 1980 and 1984. Its cross-sectional part (N=6,748) is nationally representative of adolescents born in the U.S. during that period, and the minority oversample consists of 2,236 adolescents. Information on adolescents includes basic demographics, socioeconomic status, education, employment, health and behaviors related to health, antisocial and risky behaviors, attitudes and expectations, as well as peer behaviors. At the baseline interview, extensive information is also collected on parents, including family background, education, employment, marital and fertility histories, parenting styles, health and health knowledge, and religiosity and various attitudes.

4.1 Unhealthy behavior

Drinking: For alcohol consumption, we focus on all types of hazardous drinking. The NLSY97 asked respondents whether they drank since the date of last interview, and for those who answered yes, they were asked whether they drank in the last 30 days, and their patterns of consumption. We define *hazardous drinking* in three ways, and involvement in any of them is regarded as having had dangerous drinking in the recent past. The first measure is whether a respondent has consumed alcohol in 15 or more days in the last 30 days. We construct the second measure according to WHO's AUDIT guidelines (Babor et al.): whether a male consumes 4 or more drinks per occasion and whether a female drinks 3 or more drinks per occasion. The third measure is whether one engages in binge drinking (5+ drinks per occasion) on a weekly basis. Since the NLSY97 does not ask about information based on the CAGE criteria, this is by far the most we can make out of the data, and we should bear in mind that measurement errors in these variables could generate inexact results.

4.2 Independent variables

Current marital status. The functional form and inputs available of household health production vary with marital statuses, as discussed earlier. Basically there are three categories of marital status: never married, currently married, and the formerly married. Given the rising prominence of cohabitation, we make the distinction of currently cohabiting and those who are not among the never and the formerly married. We tried three marital statuses with interaction for current cohabitation but it did not work.

We recognize the fact that marital statuses are hardly *"independent"* variables: they are subject to individual choices and hence endogeneity. Nevertheless, with the aid of fairly extensive controls on background information and individual traits, we attempt to make them as "exogenous" as possible.

Marital history. Marital history matters and we focus on three aspects of it: the number of cohabitations one has experienced, whether one gets hitched before age 18, and the duration of the first marriage. There are two reasons why we made the distinction between one and two or more cohab unions: 1) cohabitation has become the modal path to first union experience for this cohort of young Americans, and the data show it; 2) there is some evidence indicating that cohab experiences change people's outlook of and attitudes towards marriage (Axinn & Thornton, 1992), and we surmise that those having accumulated two and more cohabitation unions could behave differently. In a now defunct thread of inquiry in family sociology, teenage marriage is often characterized as immature and impulsive, and our dummy for that tries to capture these personality dimensions, if any. Marital duration attempts to measure marital capital between spouses as well as the length of mutual monitoring and sanctioning, through which health is produced in household.

Demographic controls. We employ a standard set of demographic controls: age, gender, race, Census regions (North East, North Central, South, and West), urban areas, whether any parent is foreign-born, and whether languages other than English are spoken at home. For *baseline socioeconomic characteristics*, we use family net worth in 1997 to try to capture a family's permanent income trajectory. We divide it into four categories: negative, zero to median, median to the fourth quintile, and the fifth quintile and above. Household median and the fifth quintile net worth are calculated using information from Census Household Economic Studies P70-88 (Orzechowski and Sepielli, 2003), adjusted by GDP deflator obtained from the U.S. Department of Commerce. Socioeconomic status of family is also assessed by whether parents have ever participated in any welfare program.

Labor market status and experience. As discussed earlier, one's labor market status interacts with substance use: one had better shun alcohol altogether lest it will interfere with job performance, but less could be said of tobacco. We define full-time employment in 2004 to be working for 46 weeks or more, for there is a fairly big break in the number of people having worked that many weeks. We define part-time employment in 2004 to be working between 1 and 45 weeks. We experimented with other cut-off points, such as 48 or 50 weeks, and found the results qualitatively very similar. We define full-time employment in 2003 as having worked 44 weeks or more, for the cut-off point for that year seems to be 44. Other measures were created, but the results were very similar. We also define full-time employment in 2004 to be working for 1610 hours or more (1610 = 46*35), and find the results to be similar. For cumulative time worked in adolescence and since age 20, we use information on total weeks worked obtained from the 2004 survey and divide it by 48 to convert into years.

Prior unhealthy/risky/antisocial behaviors and peer behaviors. There is a great deal of continuity in human behaviors, and to properly control for the influence of an intimate union, we must take into account of prior substance use. Then-current and hazardous drinking and smoking in 1997 are defined as they are in 2004. For a small sub-sample of NLSY97, we also measure how many days in a school week they ate breakfasts, how many days they ate fruit and vegetable, and how many days they did exercises in 1997. In addition to self-reported and parent-reported general health, physical/emotional conditions, chronic medical conditions, BMI categories (under-weight, normal, overweight, and obese) we also used youths' knowledge on damages binge drinking can do to one's health, and their parents' such knowledge in certain cases. In addition, for antisocial/deviant behaviors, we include ever used marijuana and then-current marijuana users, ever run away, carried a handgun, joined a gang, committed property crimes, attacked others, and sold drugs in 1997. The reason to include these variables is that if individuals forming early marriages are characterized by impulsivity and/or a lack of orientation toward the future, then we should be able to tap this dimension of personality/attitudes by baseline antisocial/risky behaviors.

For peers behaviors, we generate dummies to assess the proportion of peers going to church, smoking, drinking, doing drugs, having sex, skipping schools, joining illegal gangs, participating in school activities, volunteering, and planning to go on to colleges.

Family background and parental characteristics. We use all four sets of measures of family structure: at ages 2, 6, 12, and in 1997. Intact family is the reference group, with dummies for blended/stepfamily, single-mother- and single-father-family, and others. Because of the small number of children from alternative family structure, we later combined them all to use it as the omitted group to show the impact of having grown up in intact families. Parental characteristics include parental marital and fertility history—whether they are continuously- or never-married, whether they gave births as teenagers or gave birth to the youth as teenagers—and whether they are pessimistic. It also includes parents' general health status and any debilitating long-term health problems in 1997. Parental education is divided into four groups, with high school as the omitted one.

Parental control and parenting style. These two sets of variables influence youths' substance use in 2004 by shaping their values, preferences, social circle and friends, and prior substance use in 1997. Parental controls include whether mothers/fathers know children's friends and their parents, children's whereabouts, and teachers and school activities. Parenting style variables consist of parents being supportive or not and whether parents are strict disciplinarians. We expect individuals from families where parents exercise stricter control during adolescence or have an authoritarian parenting style to be less likely to engage in various unhealthy and risky behaviors.

Religiosity. We experimented with three measures of religiosity. First, parents were asked whether they were very religious or not religious at all during the baseline interview. Second, again during the baseline interview, adolescents were asked to state their religious affiliation. We follow the practice of Laumann et al. (1994) in grouping religious preferences. The group of mainline Protestants (Type I Protestant) includes Methodist, Lutheran, Presbyterian, Episcopal/Anglican, United Church of Christ (or Congregationalist or Evangelical Reformed), Disciples of Christ (or the Christian Church), Reform (or Reformed Church in America/Christian Reformed Church). The fundamentalist Protestant group (Type II Protestant) includes Baptist, Holiness (Nazarene, Wesleyan, Free Methodist), Pentecostal (Assembly of God, Pentecostal Holiness), and Non-denominational Christian (Bible Church). Following Laumann et al., we also group all types of Jewish faiths (Orthodox, conservative, reform, and others) and all types of Latter Day Saints (the Mormons) with Type II Protestant for greater stability of coefficient estimates. Since the codebook does not provide further details on "other Protestant" (N=433), we treat them as "Type III Protestant". We also group individuals having no religion together (agnostic, atheist, and personal philosophy). Respondents were also asked whether they relied on religion for good values in 2002, and we tried it as the third measure of religiosity.

5. Analytic plan

We first present sample descriptive statistics and then move on to exploratory regressions. We use a linear probability model instead of probit/logit for the ease of interpretation in earlier stages of analysis. Results should be qualitatively similar.

5.1.1 Selection

To account for selection, we focus on the use of propensity scores to eliminate selection bias. At first, we briefly discuss how the selection equations are estimated and how propensity scores were constructed. We used the propensity scores mainly for matching. Although the results from propensity score matching are not ideal due to both the difficulties with the estimation and the assumptions underlying the matching methodology, our results could be taken as a starting point for more sophisticated procedures in the future. We also performed simple post-estimation diagnostics to see the extent to which our propensity scores succeeded or failed.

Controlling for selection turned out to be far more complicated than we first thought. Even in the simplest way—reduced-form estimation paying no attention to recursive nature of this problem, we are facing two interrelated treatments—marriage and the presence of biological children in households. In addition, the first type of treatment takes more than one value: never-married, cohabiting, and married. Further, the fact that both the outcome and the participation equations are discrete choices makes it impossible to apply traditional econometrics techniques.

It has been suggested to us that we could simplify our empirical works by combining types of treatment and disregarding fertility choices. We decided to implement these procedures later on in this paper for the following reasons. First, as discussed earlier, cohabitation and marriage are qualitatively very different. Although more than half of all marriages today are preceded by cohabitation (Bumpass and Lu, 2000), cohabitors are still characterized by their non-traditional, individualistic attitudes, as we found in our descriptive results (not shown). To combine cohabitation with marriage will essentially violate the monotone treatment assumption (Manski 1997). Imagine a scale of attitudinal conventionality. Suppose that people who get married, compared to the never-married, are more conventional in attitudes. Cohabitors may not be located between the two ends of the spectrum. Instead, they may be even more unconventional than the never-married people. In other words, the use of conventionality as an instrument for getting treatment may not satisfy the monotone treatment assumption. On the other hand, this problem could be cured by separating the two into different groups, conditional upon having entered into an intimate union. Second, our descriptive regression results suggest that having biological children at home is a very strong deterrent of hazardous drinking. Ignoring this variable will basically result in serious omitted variable bias, with its effects absorbed by marriage. Alternatively, coefficient estimates for marriage and cohabitation will be upwardly biased. As a result, we only acted on these suggestions later in this paper.

Traditionally, labor economists have relied heavily on instrumental variable method/two-stage least squares method to account for endogenous explanatory variables. And one may be tempted to apply the non-linear least squares method twice to both the participation and the outcome equations. However, as Wooldridge (2002, p.236, 478) pointed out, when both equations are non-linear and involve an indicator function for latent variables, inserting consistent first-stage estimates into another non-linear equation at the second stage will *not* generate consistent estimates. It is because we cannot take expectation over an indicator function, which is non-linear, too. Unless we use maximum likelihood procedures, predicted values from the choice equation cannot be plugged in the second stage. It is for exactly the same reason why the control function approach ceases to apply here.

Although the literature on selection and treatment effect generally discusses binary treatment, in recent years there are developments to extend the treatment effect literature to multi-valued and multivariate treatments. Joffe and Rosenbaum (1999) first discussed the case of ordered doses and recommended the ordinal logit procedure. Imbens (2000) and Lechner (1999, 2002) extended the propensity score method further to account for unordered treatments and derived conditions under which estimates have a causal interpretation. They used propensity scores primarily for matching. Both Imbens and Lechner suggested using multinomial procedures to construct marginal propensity scores for matching, and this is called a "structural" approach in Lechner (2002) because all alternatives, including non-participation, are considered at the same time. Lechner also suggested a "reduced-form" procedure to simplify analyses by performing a series of binary comparisons. This is made possible by invoking the conditional independence assumption (CIA), namely that the participation decision and outcomes are independent conditional on a set of observed attributes. In addition, the reduced-form approach has the benefit of being flexible because it focuses only on pair-wise comparisons. On the other hand, since these estimations are not parametrically nested, we cannot recover structural estimates from those obtained by reduced-form procedures.

Imai and Van Dyk (2004) demonstrated how Imbens's methods could be put into practice. Upon obtaining the propensity score for each treatment, which could be ordered, unordered, or continuous and there could be more than one treatment, they recommend subclassification of the data by propensity scores into equal-sized groups. Within each sub-class, individuals should be homogeneous with respect to the propensity scores, and subclassification has the benefit of being non-parametric and more robust. Having divided the data into sub-classes, one could use matching or regression to obtain treatment effects. If regression is adopted, it is recommended that all covariates and raw treatments be used. Coefficient estimates and standard errors will be weighted by cell sizes and the results are the estimated causal effect. We follow this procedure in our attempts to remove selection bias.

5.1.2 Selection/choice equations

We estimated the selection/choice equation to construct propensity scores. With respect to the choice of model, since the decision to marry/cohabit is a discrete choice model, we used multinomial probit instead of logit because the former did not impose the assumption of independent irrelevant alternatives (IIA) (Caliendo (2006), Lechner (2002)). The literature indicates that using the marginal probabilities directly from multinomial probit estimation is allowable, although Lechner and Lee (2005) also suggested using conditional probabilities based solely on the two alternatives considered as the propensity scores in pair-wise comparisons.

Nevertheless, in terms of variable choice, there is no consensus thus far on which variables should be included. Imai and Van Dyk (2004) recommended using all covariates. Caliendo (2006), however, suggested using only the variables that can simultaneously influence choices and outcomes, and variables that can recursively affect the participation decision should be excluded. Caliendo further pointed out that controlling for irrelevant variables may result in problems of common support and higher variance of estimates. Since the purpose of a propensity score is to reduce high-dimensional matching to one (or only a few) and to balance the covariates, variables related to participation but not outcomes can be excluded, as indicated by Lechner (2002).

6. Descriptive Statistics

We present weighted sample descriptive statistics in the following table:

Table 1: Summary Statistics for the NLSY97 Analytic Sample in 2004 (weight = 04)

	All	Never Married Not Cohab	Never Married Cohabiting	Married	Disrupted Not Cohab	Disrupted Cohab	Remarried
Age04	22.021	21.821	22.226	22.617	22.712	22.699	23.649
Male	0.505	0.559	0.404	0.385	0.474	0.146	0.184
White	0.688	0.669	0.708	0.752	0.672	0.831	0.818
Black	0.143	0.166	0.126	0.063	0.077	0.032	0.070
Hispanic	0.125	0.115	0.136	0.151	0.214	0.137	0.022
ParForeign Born	0.079	0.084	0.058	0.077	0.096	0.025	0.000
South	0.360	0.346	0.324	0.445	0.549	0.415	0.493
West	0.210	0.211	0.225	0.194	0.125	0.191	0.234
HHNetW5080	0.300	0.305	0.292	0.290	0.220	0.197	0.400
HHNW5thQnt	0.214	0.240	0.145	0.179	0.166	0.130	0.168
Current Drinker	0.693	0.715	0.685	0.584	0.701	0.859	0.701
All Hazard Drink	0.499	0.534	0.488	0.349	0.509	0.490	0.344
N	4685	3257	729	598	49	27	13

The average age of the analytic sample is 22 years old, with almost one-half of them are males. There are 14.3% African Americans and 12.5% Latinos. Married people are slightly older than the never married but slightly younger than the disrupted and the remarried. There are also many more females among the married and the formerly married. This is likely because males in 2004 were still too young to marry and to support a family. Among the married and some of the formerly married, more are whites, while blacks concentrate heavily on the never-married. In addition, more married people are from the South than elsewhere.

For family net worth in 1997, more of the married and the never-married-not-cohabiting people came from families with net worth between the median and the 4th quintile range, and more of the never-married, not cohabiting came from families with net worth above the 5th quintile in 1997.

The last 2 rows show the proportion of respondents within each marital group engaging in drinking and heavy drinking. The married indeed exhibit lower propensity to engage in each of them. We use descriptive regression to clarify the relationships. Table 2 shows detailed results from descriptive regression of unhealthy or risky behaviors on marital statuses. We should bear in mind that they are regressions of one endogenous variable on another, and they only indicate the direction and magnitude of associative relationships instead of causal ones. Although the existence of such relationships does not mean much, their absence would certainly prevent us from proceeding further.

Table 2: Descriptive Regressions of Unhealthy Behaviors on Marital Statuses

Unhealthy/ Risky Behaviors	N	Never Married	Disrupted	NevMar NotCoh	NevMar Cohab	Disrupted NotCoh	Disrupted Cohab
HzDrk1	5250	0.073***	0.024	0.078***	0.050***	0.024	0.025
HzDrk2	4606	0.110***	0.118**	0.115***	0.087***	0.207***	-0.019
HzDrk3	4555	0.129***	0.032	0.142***	0.068***	0.142**	-0.158*
HzDrkAll	4581	0.128***	0.105*	0.131***	0.112***	0.178**	-0.011
SchWkDrk	4640	0.051***	0.028	0.054***	0.036**	0.008	0.062
Binge Drinker	4555	0.150***	0.032	0.158***	0.115***	0.151**	-0.173*
Current Drinker	7386	0.113***	0.122***	0.118***	0.090***	0.104*	0.154**
HzDrk1All	7299	0.058***	0.028	0.062***	0.041***	0.024	0.036
HzDrk2All	7363	0.116***	0.123***	0.121***	0.092***	0.163***	0.054
HzDrk3All	7312	0.099***	0.039	0.109***	0.056***	0.104**	-0.083
HzDrkAAll	7238	0.137***	0.128***	0.143***	0.114***	0.163***	0.068
Binge Drinker2	7312	0.144***	0.072	0.152***	0.112***	0.138**	-0.051

*: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$ (two-tailed tests)

For measures on hazardous drinking, whether we condition them on current drinking does not seem to affect the coefficient estimates of current marital statuses. The never married are consistently more likely than the currently married to engage in hazardous or even binge drinking, but the divorced/separated are less consistently so. It did appear, though, that for measures based on established guidelines (HzDrk2 from WHO's AUDIT) we could observe large and significant difference between the disrupted and the currently married. If we split the never married into those who are currently cohabiting and those not, we that found those who are not currently cohabiting are even more likely to drink dangerously, presumably because there is no one to monitor their life on a daily basis. Similarly, if we separate the disrupted into two groups, except for the measure of being a current drinker, those who are living with a cohabiting partner have consistently lower probability of hazardous drinking, and sometimes even less likely than the currently married.

7. Analysis

7.1 Drinking – Hazardous Drinking

If drinking per se is not regarded as a "vice" or a threat to health for many, we then investigate hazardous drinking behaviors. By hazardous drinking we meant that any of the following three conditions are met: 1) drinking in 15 or more days of the past 30 days; 2) 4 drinks or more per occasion of drinking for males and 3 drinks or more per occasion for females; and 3) ever binge drinking (5+ drinks per occasion) in the last 30 days. The gap in hazardous drinking among different marital groups is similar: except for the disrupted and currently cohabiting, a significantly higher proportion of the un-married engaged in hazardous drinking behaviors and the magnitude is substantial: greater than 10%.

In column 2 of Table 3 we report marital history. The number of cohabitation is significantly and positively associated with all hazardous drinking. The powerful discouraging effect of the presence of biological children in household remains and is still about 20%. The duration of first marriage flipped sign and significance in various specification, and we caution against over-emphasizing its effect. Early marriage itself is always insignificant.

The results after controlling for demographics are shown in column 3. Being male, older, living in an urban area, and coming from families with above median net worth are all positively related to getting dangerously drunk, while being black and other racial/ethnic minorities, having a foreign-born parent, and living in the South are all negatively associated with hazardous drinking last month. It should be noted that the coefficient for males is twice as large as in the case of current drinking.

Column 4 reports results controlling for marital history and demographics. Note that the strong effect of presence of biological children at home is cut in half, yet our previous conjecture that age confounds the effects of the disrupted cohabiting is unfounded. There is simply no effect for this group at all. Coefficient estimates for other marital and demographic groups are largely unchanged.

Inclusion of enrollment and employment in column 5 raises the marital status somewhat, with marital history and demographics unchanged. Labor market status and history are in unexpected sign: working full- and part-time are both associated with higher probability of dangerous drinking, and so is cumulative time worked in adolescence. We really suspect whether workplace socialization opens the door to drinking for these young people. We further controlled for current educational attainment in column 6, and found nothing significant.

For prior unhealthy behaviors, we first controlled for knowledge on damages of binge drinking. Model fit increases to about 0.15, yet sample size shrinks rapidly to just under 1,000. Consequently we do not control for prior health knowledge. For prior unhealthy behaviors per se, we experimented several specifications, and found the combination of ever drinking and any hazardous drinking in 1997, and ever smoked in 1997 to be the most satisfactory. In addition, we found that once we controlled for more than 2 measures of prior drinking, the numbers of cohabitation ceased to matter, and so did age. We also experimented with youths' prior antisocial/deviant behaviors, as well as peers drinking and smoking in 1997, and peers' antisocial/deviant activities, and found none of these reached statistical significance.

In column 8 we report results after accounting for the structure of family of origin. All 4 sets of measures (taken at ages 2, 6, 12, and in 1997) yielded unexpected outcomes: those from non-intact families are *less* likely to engage in any hazardous drinking, and sometimes significantly so. When combined with prior unhealthy behaviors, the same pattern persists. Therefore we probably could say that family structure does appear to have impacts on youths' hazardous drinking.

Again we experimented with parental characteristics and marital history in column 9. The family structure under which parents were raised is never significant, so are becoming parents as teenagers and parental pessimism. Parents' being married continuously is in the expected direction and also significant. Maternal and paternal education is positively associated with hazardous drinking, as in the case of current drinking. Parental religiosity is strongly negatively associated with such behaviors, with magnitudes ranging from 6% to almost 10%. Rather unexpectedly, once we take it into account, the significance of current marital status is drastically attenuated.

Finally we tried parental control and parenting styles. Most of the parental controls appear to discourage hazardous drinking, albeit insignificantly so. Whether parents are supportive or strict in discipline do not appear to have important effects. None of the coefficients of marital status are attenuated as a result of their inclusion.

As a result, for the final regression of hazardous drinking, we would include a full set of human capital controls, prior unhealthy behaviors, family structure in 1997, parental marital history and religiosity, and maternal education, besides standard marital status and history and demographics.

7.2 Final Results from Descriptive Regressions

Drinking: There are two patterns on alcohol consumption. First, drinking is positively associated with higher socioeconomic statuses: family net worth, parental and own education, and own employment status. The last two seem to contradict our theory that people will drink less in consideration of working. The positive signs of employment status and history instead appear to suggest that employment opens up doors to business socialization, which results in more drinking. This is consistent with previous findings summarized in Cook and Moore (2000) and Grant and Litvak (1998).

Second, current marital statuses have powerful dampening effect on drinking. Further, the formerly married are much more likely to drink than the never-married, indicating the perils of marital failure found elsewhere in the literature and that marriage is indeed a risky business. Still, the negative impact of marriage on alcohol consumption sustains whatever controls are added and only attenuated somewhat. The presence of biological children in household also has significant depressing influences on both measures of drinking and, unlike in the case of smoking, it persists in every specification.

7.3 The construction of propensity scores

As discussed earlier, there is no consensus on how to construct propensity scores in the literature yet. We therefore started by using the variables based on their correlation with the marriage decisions and heavy drinking behaviors. Table 6 shows the variables included in the marriage equation, and these are basic demographics (sex, age, race/ethnicity), geographic location, ever used marijuana in 1997, peers' substance use in 1997, enrollment status and educational attainment in 2004, and cumulative time of employment since age 20. For the fertility equation, we included basic demographics, family socioeconomic status in 1997 measured by net worth, ever used marijuana or ran away in 1997, peers behaviors, enrollment status and educational attainment in 2004, cumulative time worked in teenage years, family structure, youths' and parents' health, and maternal education. We call this model 1. We used multinomial probit to estimate marginal probabilities as propensity scores.

To assess the quality of this set of propensity scores, we then performed a diagnostic recommended in Imai and Van Dyk (2004). We first regressed all 78 covariates on each treatment in three separate sets of regressions. For each treatment, we collected all 78 t-statistics for the coefficient estimate of treatment to form a single series. Afterwards, we again regressed all 78 covariates on each treatment plus its own propensity scores (marginal probabilities). Again, for each type of treatment, we obtained all 78 t-statistics for coefficient estimate of treatment in one single series. In the end, we have 6 series of t-statistics, two (with or without own propensity score as a regressor) for each treatment (2 marital statuses and one fertility). We then plot these t-statistics on normal quantile plots and obtain 6 such plots. (Results of all these regressions (78*3*2 = 468) are available upon request.)

The idea behind this diagnostic is simple: if a treatment is random (it balances covariates), then the t-statistics from regressions of covariates on a treatment ought to be insignificant. And the t-statistic series in a standard normal quantile plot should be flat. On the other hand, if a treatment is not random and covariates are not balanced, then we should have t-statistics that are way off the [-2, 2] range. Similarly, if the covariates are balanced after adding own propensity score as a regressor, then we should see t-statistics falling mostly within the insignificance range [-2, 2]. In other words, comparing the standard normal quantile plots before and after the addition of own propensity score helps to evaluate the performance of a propensity score. Although this procedure appeared to be tedious, we should note that it is no more than running a series of Hausman tests of endogeneity on treatment variables, where the presence of endogeneity is detected by the significance of the main effect after controlling for both the main effect and its fitted value.

Figures 1 & 2 are the normal quantile plots of the t-statistics for the coefficient estimate of cohabitation and Figure 2 has the propensity score of cohabitation as an extra regressor. Similarly, Figures 3 & 4 are the normal quantile plots of t-statistics for coefficient estimate of marriage, and Figure 4 has propensity score of marriage as an additional regressor. Figures 5 & 6 are the normal quantile plots of t-statistics for coefficient estimate of childbearing, and Figure 6 has propensity score of childbirth as another regressor.

It is apparent from Figures 1, 3, and 5 that the treatments hardly balance the covariates. A large number of the t-statistics fall outside the range of [-2, 2], and some have very large values. Adding own propensity scores substantially improves the balancing of covariates: many more t-statistics are now within the range [-2, 2], as shown in Figures 2, 4, and 6. Further, even if some of the t-statistics still fall outside [-2, 2], their magnitudes are substantially reduced. In this sense, we could conclude that the propensity scores we generated are reasonable and acceptable, though there is still some room to improve to balance more covariates, which we will turn to next.

Since there were covariates that still did not balance after accounting for marginal probabilities from model 1, we added other variables in the participation equations in attempts to generate propensity scores that balance better. Ideally, the more variables affecting both participation and outcomes accounted for, the better performing the propensity scores are. Yet, in practice a large number of covariates imply a fast-shrinking sample size, and a small sample would cast doubts on the representative-ness of our results and our ability to make inference based on them.

In the end, we settled upon what we called model 2. In addition to variables from model 1, we added family socioeconomic status in 1997 measured by net worth, cumulative time worked in teenage years, employment status, intact family in 1997, maternal education, and parental religiosity to the marriage equation. There is not much change for the fertility equation.

Again we checked how well the propensity scores from model 2 balanced the covariates. Many of the remaining unbalanced covariates from model 1 were balanced in model 2, albeit at the cost of substantially reduced sample size. The one covariate that still had a t-statistics greater than 2 after accounting for model 2 propensity scores is parental religiosity, which has a higher correlation with outcome variable than many other covariates. As a result, we would include it along with additional endogenous variables in Mahalanobis matching. Further, since the sample size of model 2 is only about one-third of the original sample, as shown in Table 9, we retained our model 1 propensity scores for matching that requires a larger sample size, such as Imai and Van Dyk's subclassification method later.

Figure 1: Normal quantile plot of t-statistics of regressions on cohabitation

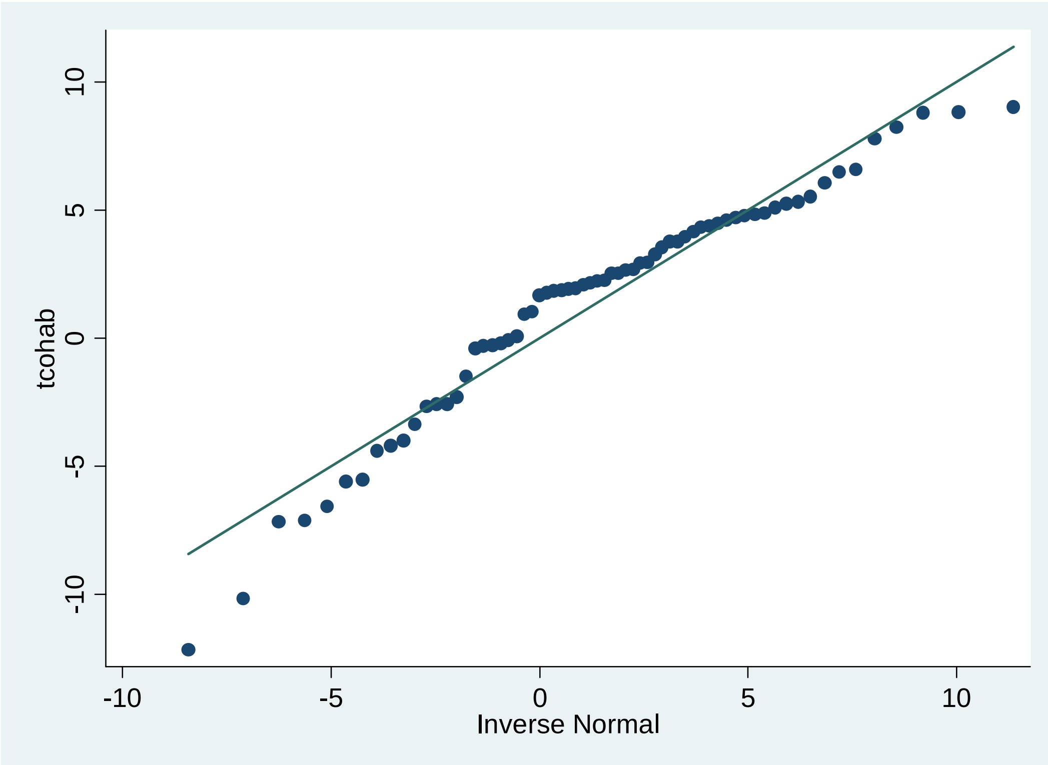


Figure 2: Normal quantile plot of t-statistics of regressions on cohabitation and its own marginal probability

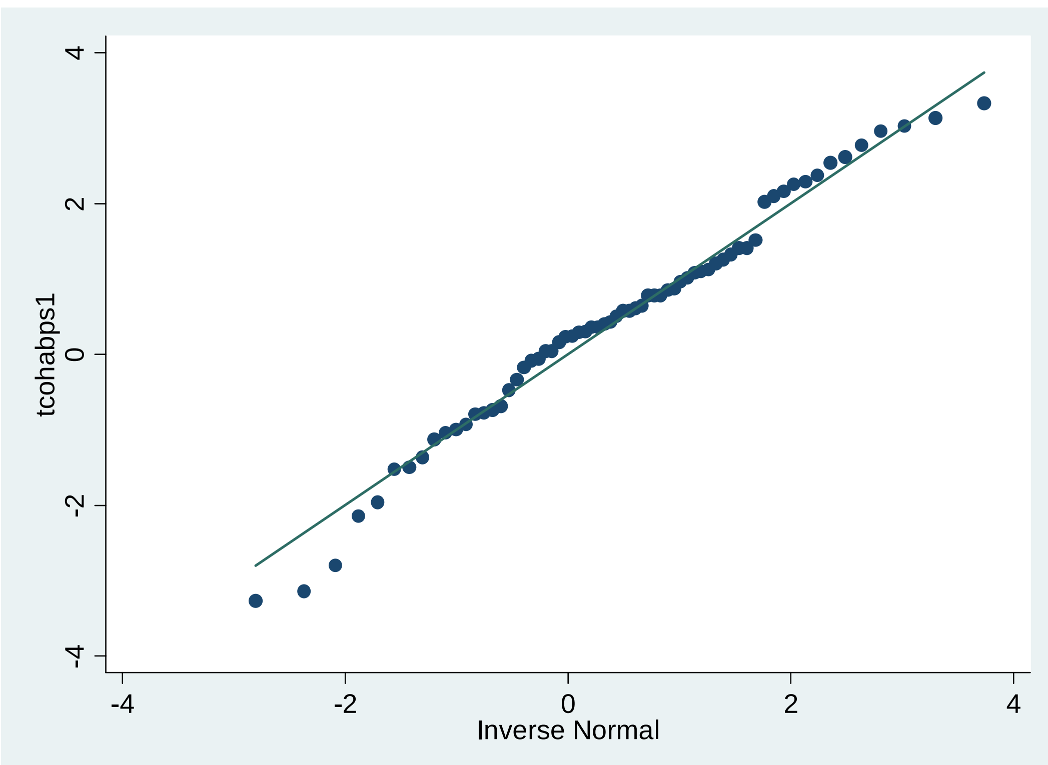


Figure 3: Normal quantile plot of t-statistics of regressions on marriage

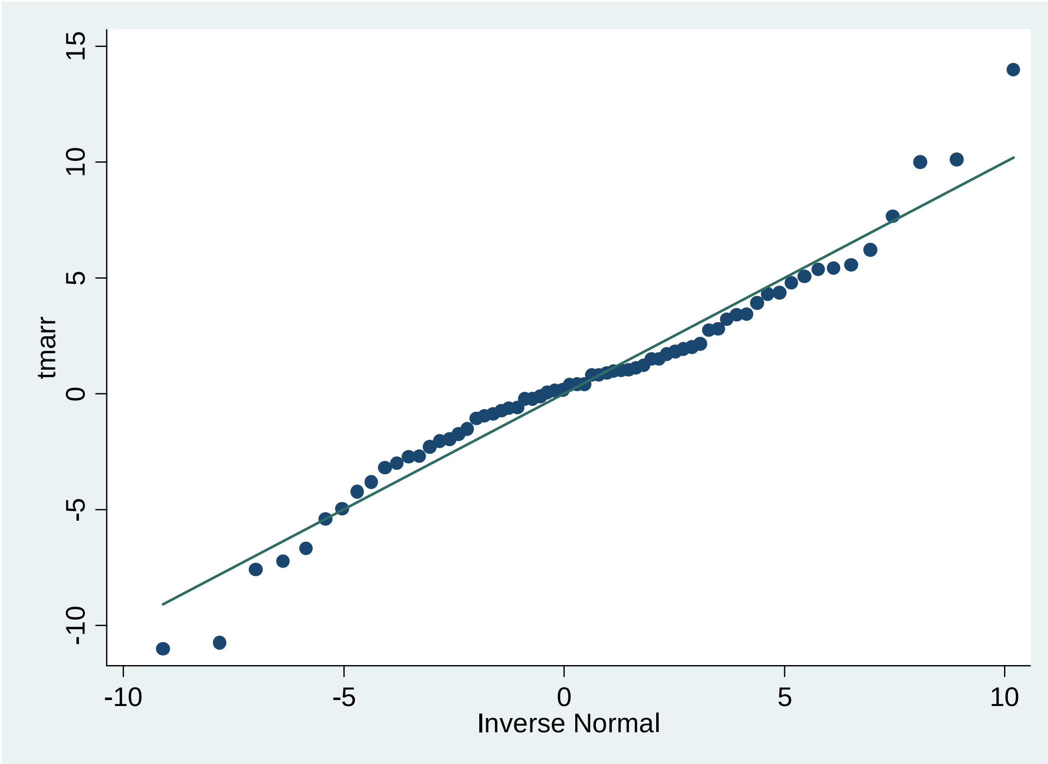


Figure 4: Normal quantile plot of t-statistics of regressions on marriage and its own marginal probability

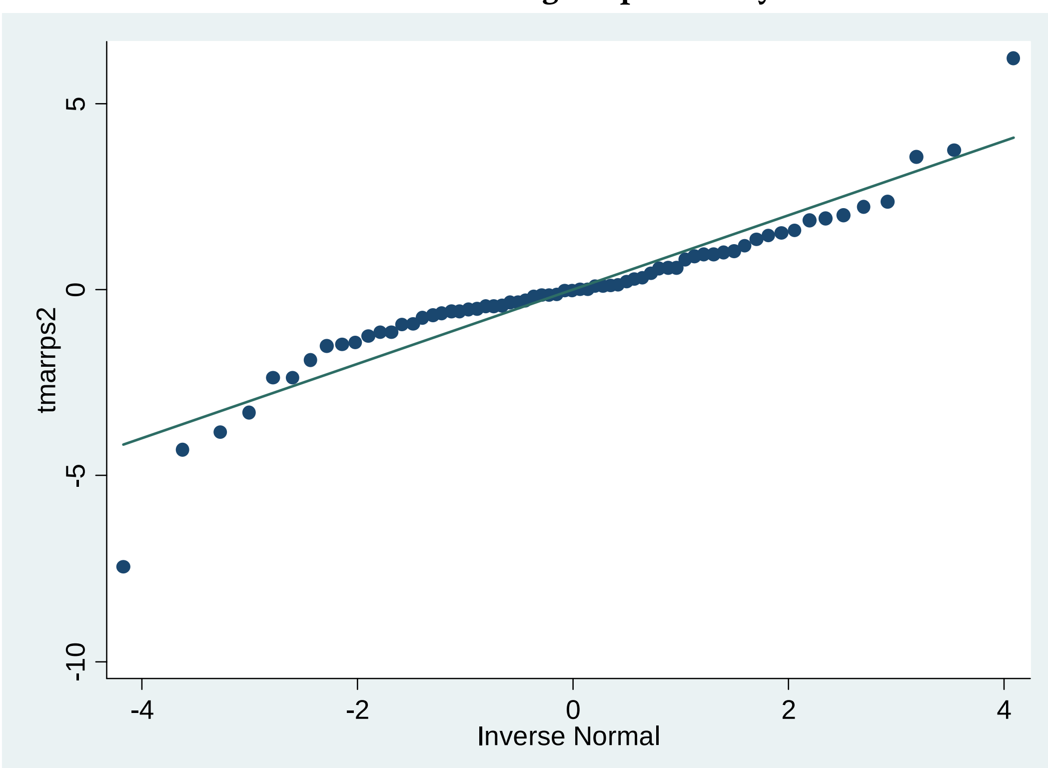


Figure 5: Normal quantile plot of t-statistics of regressions on childbearing

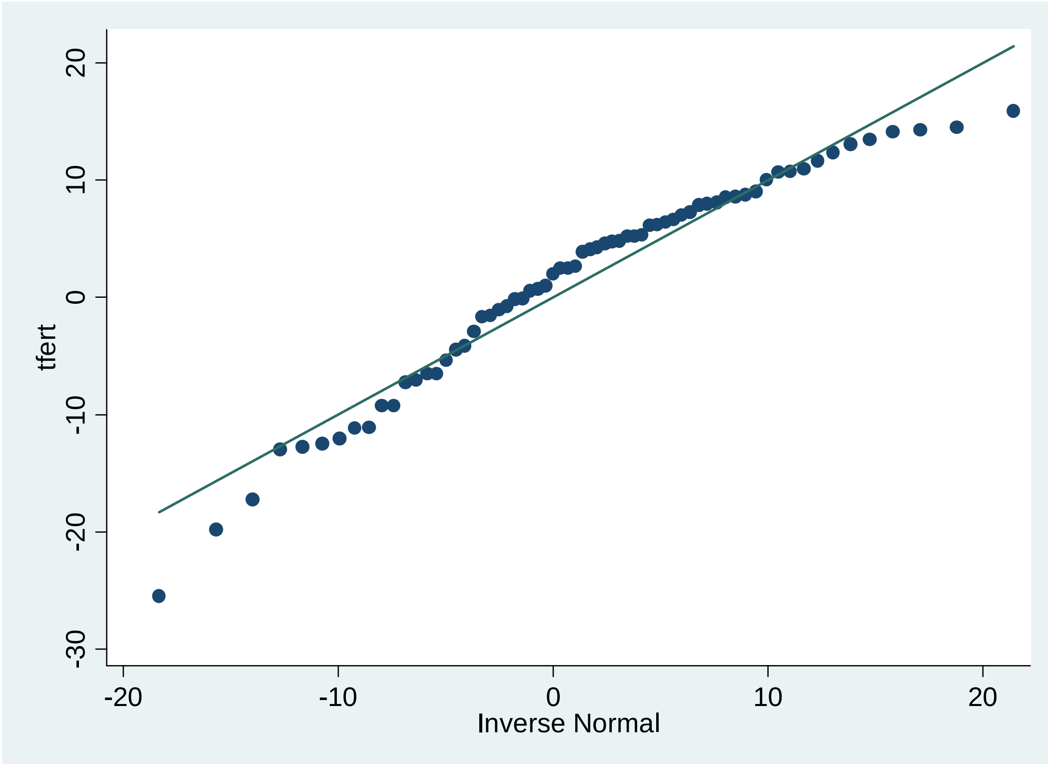
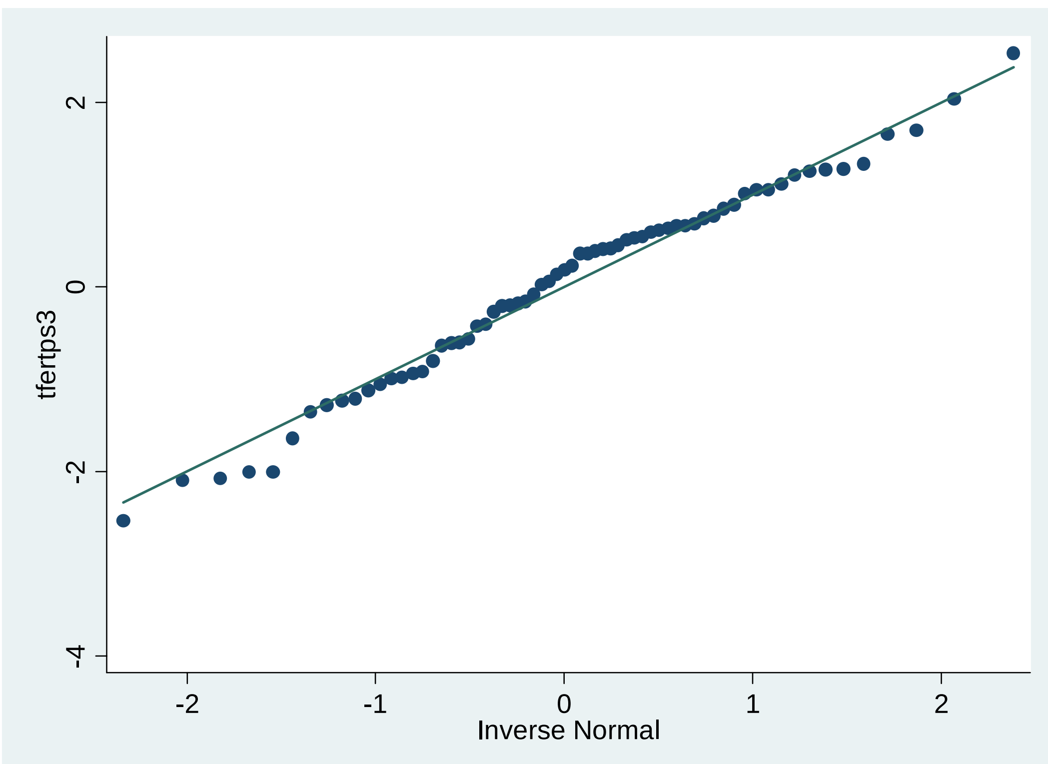


Figure 6: Normal quantile plot of t-statistics of regressions on childbearing and its own marginal probability



7.4 A brief note on sampling and weights

Since we do not have access to restricted-use geocode data of NLSY97, and the strata data are not released in order to protect the privacy of respondents, we can only use the oversampling indicator as the stratum variable. We set the primary sampling units (PSU) to the individual, employed the weights in 2004, and restricted the analytic sample to those who have not divorced or remarried in 2004.

We also experimented with other procedures. Again we restricted the analytic sample to those who have not divorced/remarried in 2004, clustering around household ID to take care of heterogeneity due to same family, and ran robust regressions with weights equal to that of 2004. In the end we found the results from both procedures discussed here are identical down to the last digit.

8. Results

8.1 Exploratory analysis

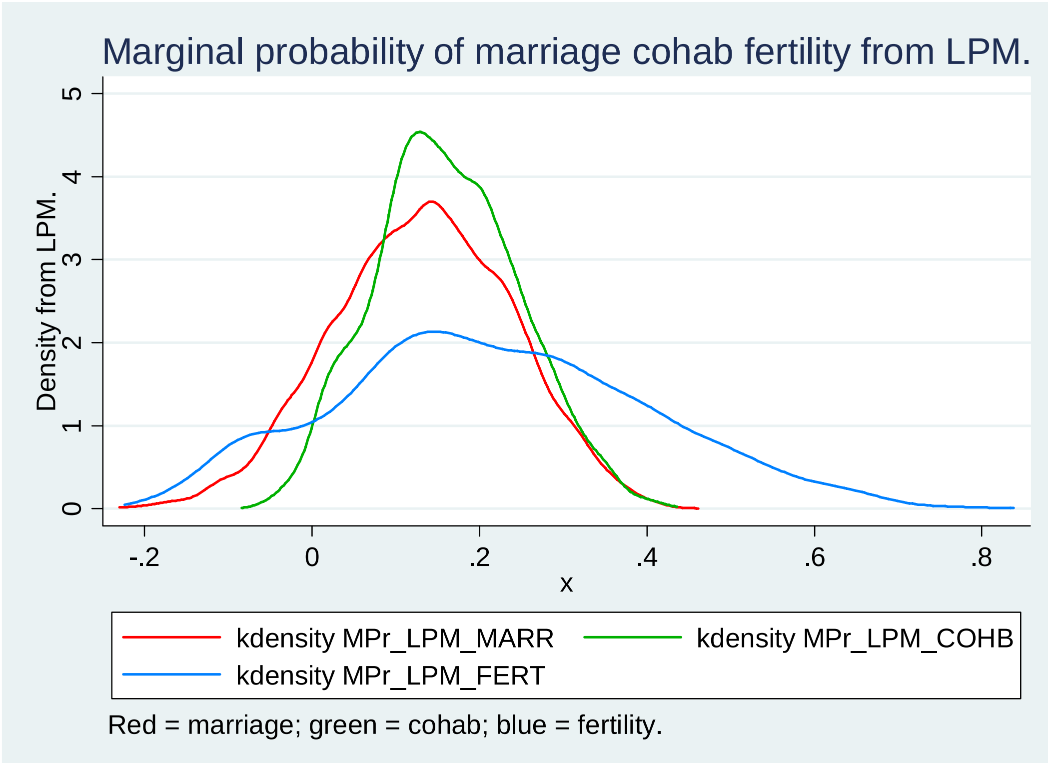
For expository purposes, we first performed a two stage least squares (2SLS) for both the participation and the outcome equations, keeping in mind the various caveats of the linear probability model. To circumvent the issue of heteroskedasticity, a major problem with OLS applied to discrete choice models, we used the Huber-White estimator.

Table 8 indicates that, when we performed 2SLS and used the fitted values for both union statuses and parental status in the equation for heavy drinking, coefficient estimates became much attenuated and lost their significance, although still with anticipated signs. Marriage is related to an 11.6% reduction in heavy drinking, while cohabitation is associated with a 12.4% increase. Having a child at home is related to a 17,8% decrease in heavy drinking. However, none of these effects were significant.

It appears that a very crude correction for endogeneity may have removed most selection bias from earlier results. Except for the following few, coefficient estimates of most covariates are also insignificant. Being a male is positively related to heavy drinking (0.090), and so are coming from families of higher socioeconomic status (0.076), being currently employed regardless of full-time (0.155) or part-time (0.133), and parents' not being very religious (0.064). On the other hand, youths' having a more conservative religion (type 2 Protestants, all types of Jewish faiths and Latter Day Saints) is negatively associated with heavy drinking (-0.132). Substance use also exhibited a very strong path dependency: having ever consumed alcohol back in 1997 is associated with a 12.1% increase in heavy drinking 7 years later, and having ever smoked in 1997 is also related to a 9.3% increase in heavy drinking, revealing a complementary relationship between drinking and smoking.

On the other hand, results from the linear probability model may have problems. In particular, the majority of our sample was still in their early twenties and naturally has never been married, and this is reflected in the propensity scores we generated. Figure 7 reveals that the great majority of the three propensity scores is below 0.2, and a non-trivial part of them is below zero. Since OLS weighs each case equally, it is possible that our results may be driven by these non-participants. As a result, we turned to other methods which could take better care of this problem.

Figure 7: Marginal probabilities of marriage, cohab, & fertility from LPM.



8.2 Propensity score matching – Mahalanobis matching with propensity scores

We started with Mahalanobis matching using propensity scores alone from model 1. We call this model 1a. Mahalanobis matching is used instead of propensity score for three reasons. First, it allows for controlling for all three propensity scores simultaneously. Second, it allows for accounting additionally for variables that still failed to balance after propensity scores. Third, it also allows us to control for other endogenous variables such as respondents' cohabitation history. On the other hand, matching on a higher number of dimensions requires a large sample size. As a result, Mahalanobis matching would not be used when many small cells are involved. Results from Table 9 suggested that marriage was associated with a 13% to 15% reduction in hazardous drinking in 2004, and these effects were highly significant. Even after controlling additionally for family background, prior (heavy) drinking and smoking, employment histories, and the number of prior cohabiting unions in model 1b, Mahalanobis matching results still showed that marriage was significantly related to a sizable reduction in hazardous drinking, and its size remained similar. On the other hand, in both models 1a and 1b, cohabitation was not related to any reduction in hazardous drinking.

Since the propensity scores from model 1 left some covariates unbalanced, as shown in Figures 1, 3, & 5, we also used the adjusted propensity scores from model 2, at first without further controlling for endogenous variables. Again, results from Mahalanobis matching revealed a 9% reduction in heavy drinking associated with marriage, and the effect was significant. Model 2b included (heavy) drinking and smoking in 1997, parental religiosity, and the number of previous cohabiting unions. Results indicated an almost 10% reduction in hazardous drinking related to marriage. With respect to cohabitation, neither model 2a nor 2b provided any support that cohabitors are less likely to engage in heavy drinking. In other words, at least in terms of heavy drinking, cohabitors behaved like the never-married young adults.

8.3 Common support

To assess the matching quality, we first need to know the extent of common support our propensity scores have. The lack of common support can pose serious problems for propensity score matching. Here we use a simple method recommended by Caliendo (2006, p.80) to detect whether there is sufficient common support. We plotted the density of propensity scores (marginal probabilities) by treatment groups, limited to marriage and cohabitation in this case. We omitted several combinations of propensity scores and treatment groups for the sake of brevity. The omitted graphs are similar to those presented here. Figure 8 shows the density of the marginal probability of marriage for both the married and unmarried, with the latter peaked at around 0.05. Figure 9 shows the marginal probabilities of entry into a cohabiting union for the cohabitors and others. Figure 10 gives the marginal probabilities of becoming a parent for married and unmarried people, and the latter also peaked at about 0.05. All these three graphs are based on model 1. Figure 11 is the propensity scores of marriage for the married and unmarried, based on model 2.

It is clear that the overlap in marginal probabilities is fairly substantial for model 1, regardless whether the treatment is marriage or cohabitation. The same also applied to fertility. On the other hand, for model 2, there is a non-trivial part of the densities which failed to overlap, although things are much better for the marginal probabilities of cohabitation and fertility (not shown). The problem with our second model is consistent with what we will show next: without further controlling for several endogenous variables, Mahalanobis matching using marginal probabilities from model 2 alone resulted in a substantial increase in both mean and median standardized bias.

Figure 8: Marginal probabilities of marriage for the married and the unmarried.

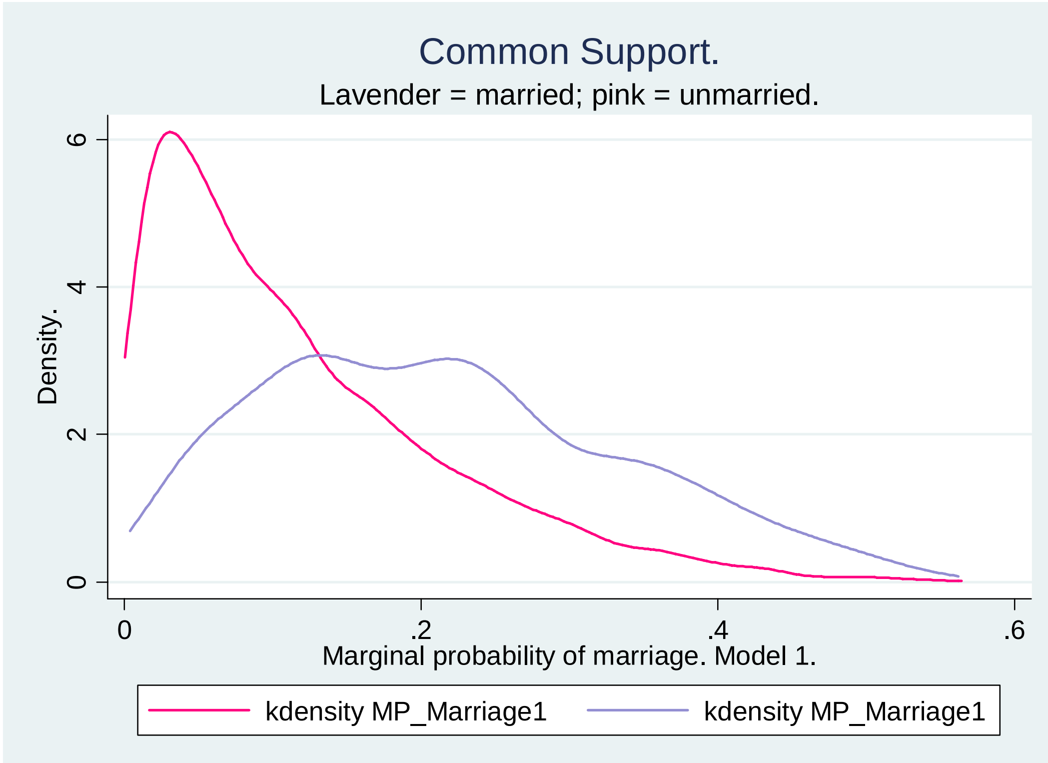


Figure 9: Marginal probabilities of cohabitation for cohabitors and others.

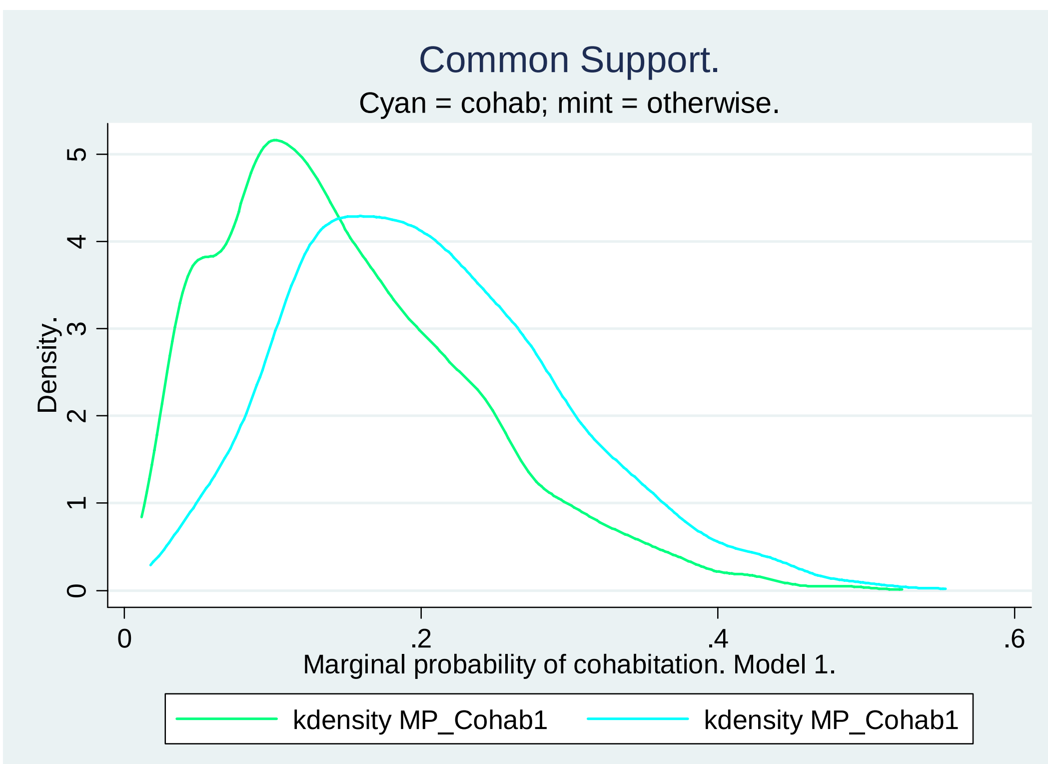


Figure 10: Marginal probabilities of becoming a parent for the married and others.

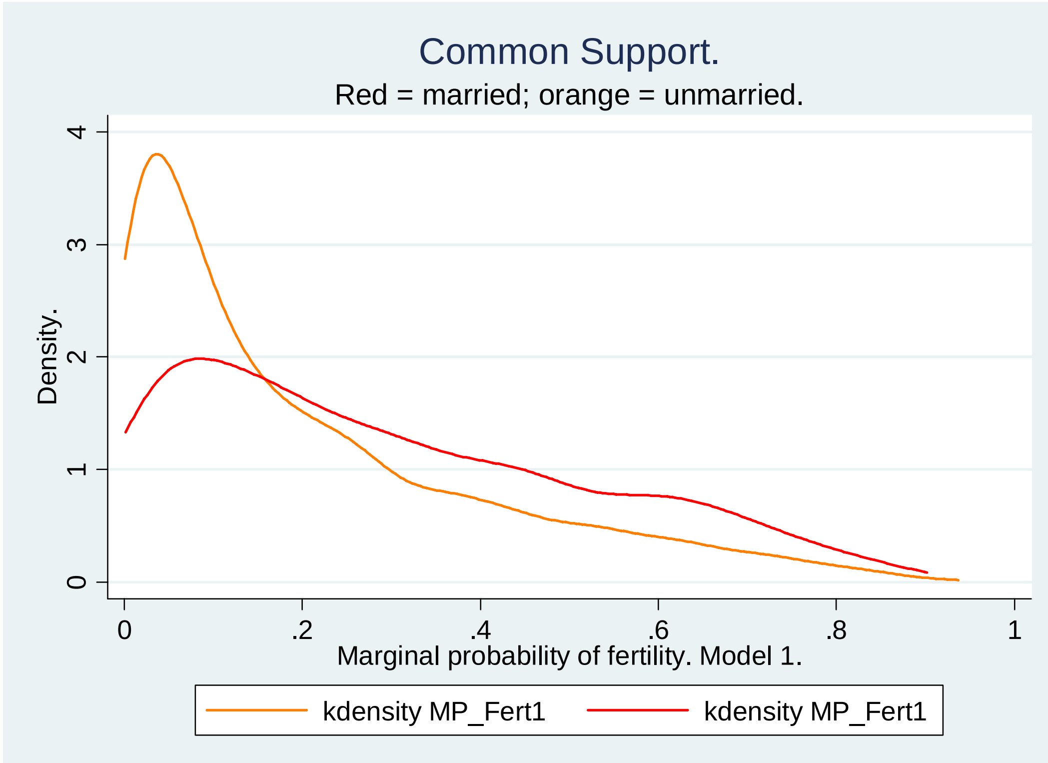
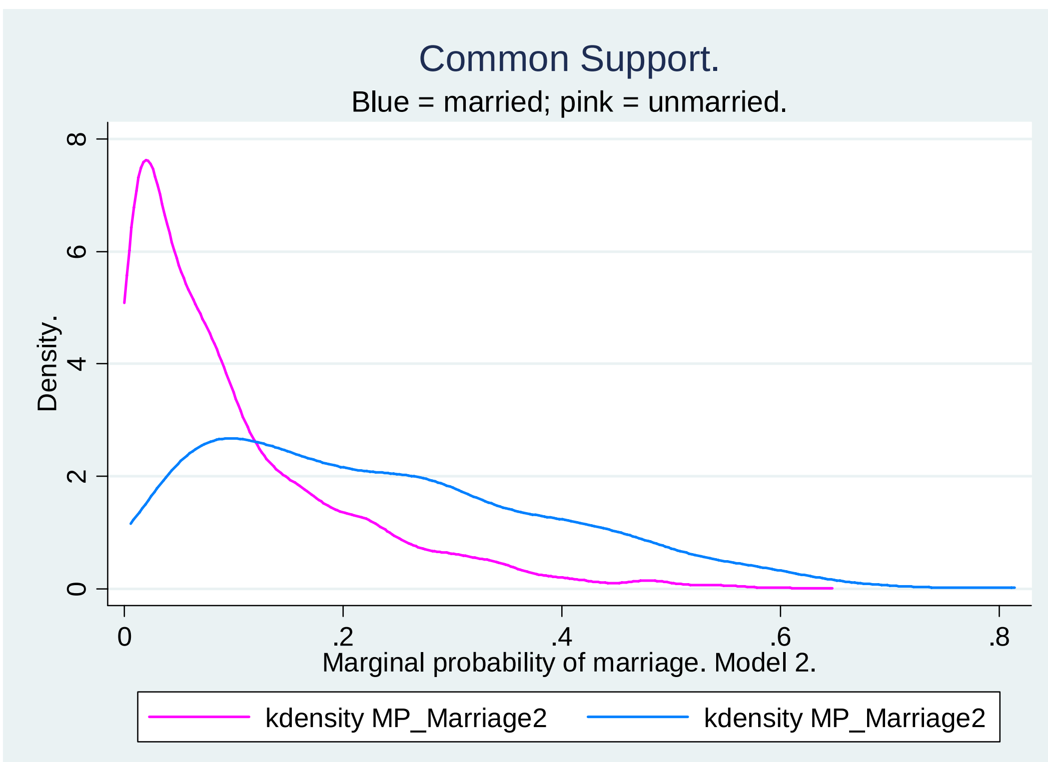


Figure 11: Marginal probabilities of marriage for the married and others, model 2.



8.4 Other measures of matching quality – Pseudo R-squared & Bias Reduction

There are other more formal ways to assess the quality of matching. Pseudo-R2 is obtained by regressing treatment probability/propensity scores on all covariates used in matching by probit and it measures the extent to which the covariates explain the participation decision (Caliendo, p.79). In an unmatched sample, this measure should be fairly high, indicating systematic differences in the distribution of covariates by treatment groups. If the propensity scores balance covariates well, then in a matched sample covariates should be similarly distributed, and pseudo-R2 ought to be low. Note that this is similar to what we have done earlier in Figures 1 to 6—a series of Hausman tests of endogeneity: if a propensity score balances well, then the treatment ceases to be an endogenous variable after controlling for propensity scores.

Table 11 shows pseudo-R2 for both marriage and cohabitation for models 1a, 1b, and 2a, 2b from Mahalanobis matching. There are clear patterns. In the absence of further accounting for endogenous variables, models 1a and 2a did not perform well according to this statistic. In fact, their pseudo-R2 *increased* after matching, and it is particularly bad for the propensity score of marriage. For cohabitation, things started out fairly well in models 1a and 2a. Adding other endogenous variables (e.g. the number of previous cohabiting unions) and variables that did not balance even after controlling for propensity scores (parental religiosity) clearly improved quality of matching substantially, particularly for the matching of marriage. In addition, adding these variables initially made the pseudo-R2 for cohabitation much larger, followed by a substantial decrease after matching.

Table 12 shows the absolute mean and median biases before and after matching. In terms of absolute biases, matching clearly removed a large part, if not the majority, of biases, and this applied to both models and to both marriage and cohabitation. However, if we standardized these biases, we found that standardized bias *increased* by no small amount for both models 1a and 2a. And the increase was particularly severe for the matching for the effect of marriage. Once we controlled for the aforementioned endogenous variables, the standardized biases decreased, but this is limited to marriage in model 1b only. For marriage in model 2b, the decrease was very small, and the increase in the matching for the effect of cohabitation there was large. This is consistent with what we found earlier that model 1b appears to be the one with the highest quality of matching.

We can take a further look at the variables of which the biases increased by more than 20% after matching. Table 13 gives these variables. It is clear that many of these ever-present variables such as youth and parental health in 1997, family socioeconomic status in 1997 measured by net worth, and parental marital and fertility histories did not enter into the estimation of marriage/cohabitation equation. They were included mainly because we wanted to control for the effect of becoming a parent on hazardous drinking. As a result, at this stage we cannot simply remove these variables in the mere hope of reducing biases after matching.

8.5 Binary matching (pair-wise comparison) using propensity scores from multinomial probit

We next turn to pair-wise comparisons (married vs. never-married people and cohabitors vs. never married) recommended by Lechner (1999, 2002). By invoking the conditional independence assumption, comparisons of several treatments can be reduced to a series of binary/pair-wise comparisons. Since our focus is on the effect of marriage and cohabiting union status on heavy drinking, we do not report results of the effects of being a parent.

For the effects of marriage, we first based our analysis on a sample consisting only of the married and the never-married, excluding cohabitors. Nevertheless, since Lechner's method basically applies to a one-dimensional treatment with multiple mutually-exclusive alternatives, it does not explicitly allow for matching on multiple-dimensional propensity scores simultaneously. In other words, for the effects of marriage, we can only match on its own propensity score, and cannot include the propensity score of parental status. To account for the influence of being a parent on heavy drinking, we divided this sample into three by the propensity scores of fertility. In this way, we can ensure that respondents in each of these three sub-samples are homogenous in terms of parental status.

We then conducted propensity score matching on each of these subsamples, first using the marginal probability of marriage, and then using the conditional probability of marriage, namely the marginal probability of marriage divided by the sum of marginal probabilities of being married and never-married. These marginal probabilities came from our earlier multinomial probit models. We then weighted the treatment effects (ATT and ATE) of these three subsamples according to sample size and obtained the overall treatment effects. To facilitate comparison with our earlier results, we performed the entire procedures on the full sample.

We repeated the same procedures on samples consisting of cohabitors and the never-married respondents as well as the full sample to obtain the treatment effects of cohabitation.

Table 14 shows the results. For every binary comparison (married vs. never-married; cohabitation vs. never-married), there are three samples (full sample, partial sample twice). We divided every sample into three by the propensity score of being a parent, and we used propensity scores from both models 1 and 2. Consequently, we conducted 36 (=2*3*3*2) times of propensity score matching and combined them into 12 sets of results. We should also note that for all the matching estimation, the proportion of sample on support exceeded 90%, and with the exception of marriage in model 2 and cohabitation in model 2 using full sample, all other matching had the proportion of sample on support well over 95%.

Marriage again is associated with a substantial decrease in heavy drinking. Regardless whether we used the full sample or a sample of the married and never-married respondents only, regardless of whether we used marginal or conditional probability of marriage, and regardless of which model these propensity scores came from, the married people showed at least a 15% lower likelihood to engage in heavy drinking, compared to the never-married respondents. Along with our earlier results, these findings show that the effects of marriage are robust to changes in sample, the specification in the estimation of marginal probabilities, and the definitions of propensity scores. These results should also put to rest any lingering doubt that our results are driven by the way we define propensity scores. On the other hand, since these results did not account for endogenous variables as we did earlier, we should read them with caution.

Cohabitation once again is unrelated to any reduction in heavy drinking, with the sole exception of the cohabitor-never-married sample using marginal probability from model 2. In all other combinations, the effect of cohabitation is either negligibly negative or positive. Along with our earlier results, findings here confirmed that the cohabitors resembled the never-married people more than the married.

8.6 Subclassification & propensity score matching

We then tried to implement the procedures recommended in Imai and Van Dyk (2004) to subclassify data according to the three propensity scores. We first divided the sample into thirds of equal size by each propensity score . In this case, we would have 27 cells, and individuals in each cell are supposed to be homogeneous in terms of the propensity scores.

As shown in Table 9, there are a large number of missing cases in model 2 of marginal probabilities, due likely to the large number of covariates used in estimating them. To maximize the cell sizes for matching or regression, we therefore decided to use the marginal probabilities from model 1. Nevertheless, Table 15 still showed that many of these 27 cells have surprisingly small number of observations. In that case, we cannot use regression to adjust for covariates, although we may still be able to implement matching in a cautious way.

This time we used simple propensity score matching because Mahalanobis matching requires more covariates in estimation than many of those small-sized cells could sustain. As discussed earlier, we used marginal probabilities from model 1 to maximize cell sizes.

We plotted the marginal probabilities to illustrate why many of the 27 cells had very few cases. Figures 12 to 14 plotted the marginal probabilities of marriage, cohabitation, and fertility. It is clear that for this sample of young adults, the great majorities had their marginal probabilities of marriage and cohabitation very close to the origin, and this is also consistent with Figures 8 to 10 where these densities peaked early. This finding is also in line with the fact that our respondents, aged no more than 24 in 2004, were just beginning to form their intimate unions—marriage or cohabitation alike. The marginal probability of being a parent is more spread out, though. Alternatively speaking, as we move away from the origin and the main diagonal in a three-dimensional space, the cases become more and more sparse.

Due to very small cell sizes, we performed a simple propensity score matching without adjusting for covariates and other endogenous variables. The treatment effects of each cell were weighted by its size, and we obtained the overall treatment effects summing over them. In the end, there were three cells each for the matching of marriage and cohabitation that failed to produce any result, either because all observations were off support, or because there were too few cases in either the treated or control group to yield standard error. Further, the proportion of the sample on support is substantially lower than what we had earlier. This is likely because we divided the sample more finely such that a non-trivial number of observations fall around the boundaries.

Table 16 gives the treatment effects of both marriage and cohabitation. Marriage is once again associated with a 17% reduction in heavy drinking, whereas cohabitation has no impact on heavy drinking at all. However, since we did not adjust for those endogenous variables, the results should be read with caution.

Figure 12: Scatterplot of marginal probabilities – marriage and cohabitation.

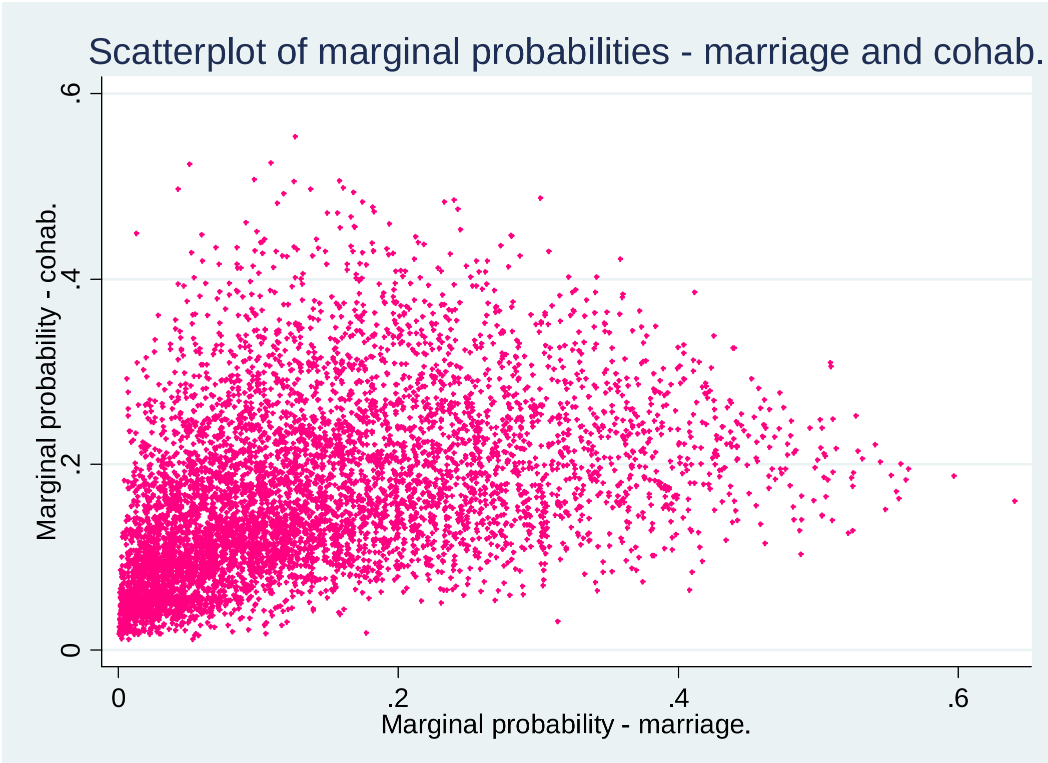


Figure 13: Scatterplot of marginal probabilities – marriage and fertility.

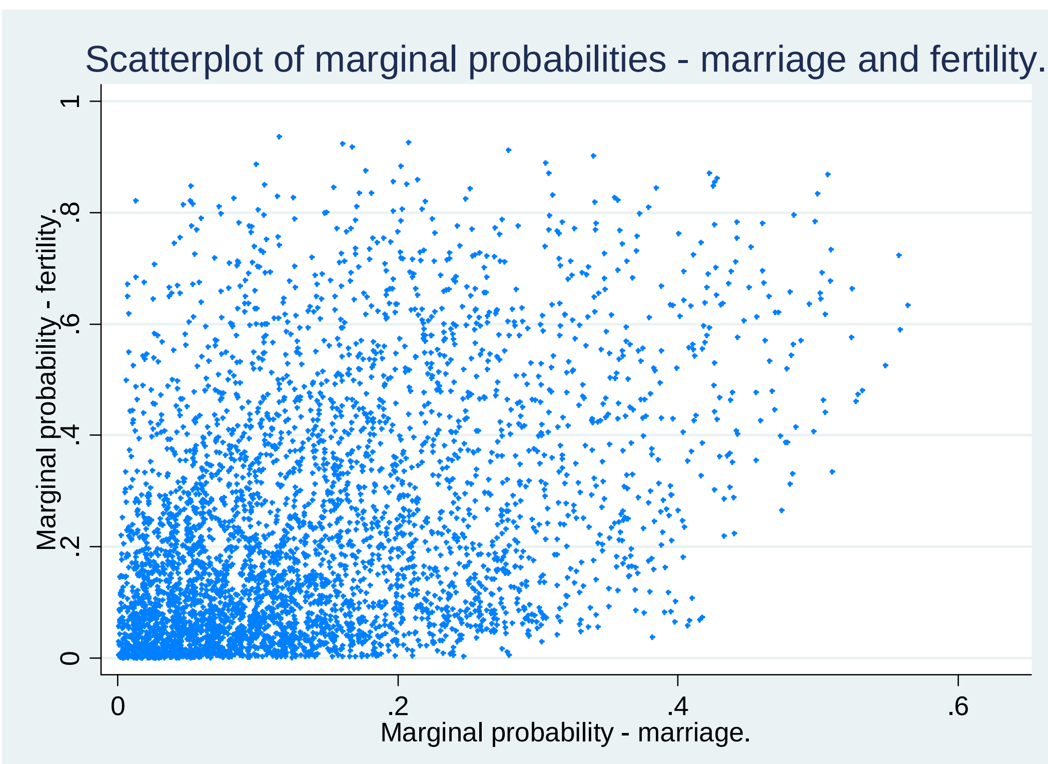
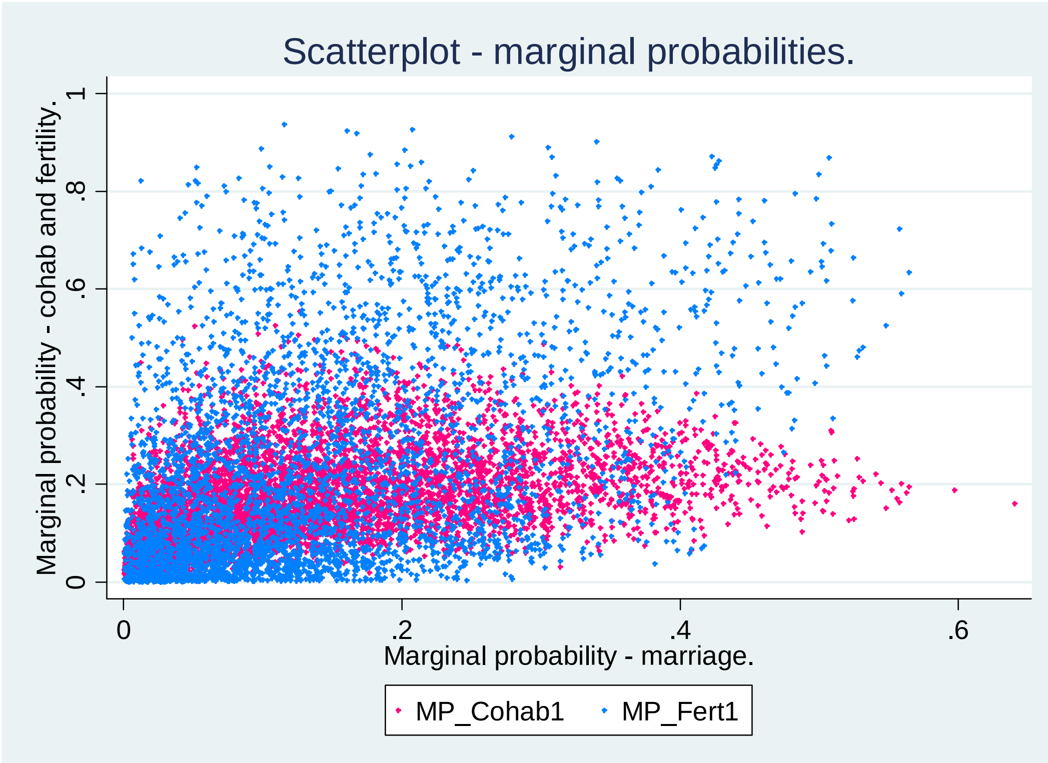


Figure 14: Overlaid scatterplots of marginal probabilities: marriage, cohab, & fertility.



In results not shown, we found that the negative impacts of marriage were not evenly spread out over the three-dimensional space. Instead, the results were driven mainly by four cells that were close to either end of the diagonal linking the origin to the end where all three marginal probabilities were the highest. Combined, these four cells had a total of 46.1% of all respondents. More interestingly, effects of marriage are the strongest in cells closer to the origin. For respondents with all three marginal probabilities in the first tercile (lowest), the average treatment effect on the treated was -0.588(!), whereas for those at the other end of the diagonal (the third tercile for all), the average treatment effect on the treated was only -0.119. This finding seemed to contradict the notion that there was sorting on the gain, where those who perceived the benefits of marriage to be the greatest were more likely to select themselves into marriage. We offer no explanation for this at this point.

For cohabitation, most results were small and insignificant throughout the 27 cells, except one significant cell where the average treatment effect on the treated was 0.294 and the average treatment effect was 0.324. They, however, were swamped by all other insignificant effects in the end.

9. Discussions

Our entire empirical works were built around matching – either Mahalanobis or propensity score matching. Matching, however, is based on the assumption of selection on the observables. If there is severe selection on the unobservables, our matching estimators may well break down. Below we offer some discussions on the nature of possible selection in this context.

Since the NLSY97 respondents were born between 1980 and 1984, those who were married in 2004 are no more than 24 years old. Earlier literature of family sociology characterized people who get married early as immature, impulsive, can not delay gratification/being less oriented towards the future, and thrill/sensation seeking. If these personality and attitudes are the unobservables that select people into early marriage and make people more likely to engage in heavy drinking, then we should be able to observe married people drink *more heavily*. Instead, for every specification used here, marriage is *negatively* related to heavy drinking. Clearly, if the aforementioned personality traits are the unobserved selection variables, their effects are negligible in this case.

It is also possible that our measure of family socioeconomic status was not good enough, and coming from a disadvantaged background selected people both to engage less in heavy drinking and to marry early. Then we could have observed that marriage is associated with *less* hazardous drinking. On the other hand, recent trends in family formation suggest that individuals from lower socioeconomic strata tend to shun formal marriage and to choose cohabitation. In this regard, it is the results of cohabitation that would have been contaminated with selection bias. Consequently, we should be able to observe cohabitors to drink less heavily than the never-married, due both to this selection effect and to the social control from the cohabiting partner. This reasoning, however, clearly contradicted our empirical results. Cohabitors drink as heavily as the never-married people, after controlling for the covariates.

The final unobservables we can think of are conservative attitudes. Parental religiosity, a proxy for an overall conservative bent, was shown to be a powerful deterrent in children's heavy drinking, second only to marital statuses and their being parents themselves. If more conservative people are more likely to marry early and to drink less heavily, then the relationships between marriage and heavy drinking will be spurious. We do not have a good measure on conservative orientation or attitudes, but we do have parental religiosity, one's own religious affiliation, and whether one needs religious beliefs for good values as proxies. Unfortunately, all of these variables only showed a strong negative correlation with heavy drinking and none showed a strong correlation with entry into marriage or a cohabiting union. In addition, none of them are strongly related to becoming a parent either. To explore this channel of selection, we need more data.

10. Conclusion

In this study we investigate the relationship between marital/intimate union statuses and unhealthy behavior—heavy drinking. We try to fill in the gap by delineating one mechanism through which health is produced in households. We situated the reduction of unhealthy behaviors within the greater context of household production of health and social control. As far as matching is concerned, our empirical works are reasonably successful, but our earlier conjecture that married people tend to disengage from heavy drinking because of concerns of work was flatly rejected by the fact that employed people were more likely to drink heavily. We tried to purge selection bias by matching, and the negative relationship between marriage and heavy drinking persisted across different samples, estimation schemes, and the definitions of propensity scores. This is not to say that our results are impeccable, but they do serve as a point of departure reasonably well.

In addition, initially we expected to find a smaller effect of cohabitation on heavy drinking than marriage, whereas we ended up with finding nothing for cohabitation, and this non-finding also persisted throughout our analysis. This suggests that, among the NLSY97 respondents, cohabitation resembles singleness more than marriage.

In the future, we will make the following improvements. First, we will adopt nested procedures for both hazardous drinking and entry into an intimate union. In other words, we will model whether an individual drinks first. Conditional on that, we then will estimate whether one drinks excessively. Similar procedures shall be applied to whether one enters into an intimate union and then whether one chooses cohabitation or formal marriage. Second, our construction of propensity scores may have imposed too much structure on the data. Later we shall use other non-parametric methods to better exploit the data. Third, here we limited ourselves to reduced form estimation and ignored possible interaction among unhealthy behaviors and entry into marriage and cohabitation.

In modern societies, marriage appears to be a beleaguered social institution that faces unprecedented challenges. Nevertheless, the difficulties marriage encounters today arise more from the public's misunderstanding than its being outdated. The fact that marriage is still related to elevated physical and mental health and likely has played a role in their emergence demonstrates that marriage is still highly relevant to individuals' well-being. The persistence of marital benefits illustrates that marriage is still able to weave together social support, social control, and economic forces to confer individuals benefits and enhance dyadic solidarity. Even in a society characterized by fragile interpersonal relations, dominating individualism, and skyrocketing divorce rates like the U.S., findings from these young adults, who came of age only recently, deliver affirmative evidence and a reassuring message that marriage still keeps up with time and is still relevant, now more than ever. In short, marriage still matters.

Muh-Chung Lin

References

Akerlof, George A. 1998. "Men without Children." *Economic Journal* 108: 287-309.

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

Babor, Thomas F., John C. Higgins-Biddle, John B. Saunders, and Maristela G. Monteiro. *AUDIT: the Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care.* Second edition. World Health Organization: Department of Mental Health and Substance Dependence.

Becker, Gary S. 1991. *Treatise on the Family*. Cambridge, MA: Harvard University Press.

Berkman, Lisa F., Thomas Glass, Ian Brissette, and Teresa E. Seeman. 2000. "From Social Integration to Health: Durkheim in the New Millennium." *Social Science and Medicine* 51(6): 843-857.

Bumpass, Larry L. and Hsien-Hen Lu. 2000. "Trends in Cohabitation and Implications for Children's Family Contexts in the United States." *Population Studies* 54(1): 29-41.

Caliendo, Marco. 2006. *Microeconometric Evaluation of Labor Market Policies*. Berlin: Springer-Verlag.

Cook, Philip J. and Michael J. Moore. 2000. Chapter 30: Alcohol. In *Handbook of Health Economics*, Volume 1B, eds. A. J. Culyer and J. P. Newhouse. Elsevier Science, B.V.

Dee, Thomas S. 2006. "Forsaking All Others? The Effects of "Gay Marriage" on Risky Sex." *NBER Working Paper* #11327.

Duncan, Greg J., 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.

Franks, Melissa M., Amy Mehraban Pienta, and Linda A. Wray. 2002. "It Takes Two: Marriage and Smoking Cessation in the Middle Years." *Journal of Aging and Health* 14(3): 336-354.

Ginther, Donna K. and Robert A. Pollak. 2004. "Family Structure and Children's Educational Outcomes: Blended Families, Stylized Facts, and Descriptive Regressions." *Demography* 41(4): 671-696.

Goldman, Noreen. 1993. "The Perils of Single Life in Contemporary Japan." *Journal of Marriage and the Family* 55(1): 191-204.

Goldman, Noreen. 1993. "Marriage Selection and Mortality Patterns: Inferences and Fallacies." *Demography* 30(2): 189-209.

Goldman, Noreen. 2001. "Social Inequalities in Health: Disentangling the Underlying Mechanisms." *Annals of the New York Academy of Sciences* 954: 118-139.

Grant, Marcus and Jorge Litvak, eds. 1998. *Drinking Patterns and their Consequences*. Washington D.C.: Taylor & Francis.

Hu, Yuanreng and Noreen Goldman. 1990. "Mortality Differentials by Marital Status: An International Comparison." *Demography* 27(2): 233-250.

Imai, Kosuke and David A. Van Dyk. 2004. "Causal Inference With General Treatment Regimes: Generalizing the Propensity Score." *Journal of the American Statistical Association* 99(467): 854-866.

Imbens, Guido W. 2000. "The Role of the Propensity Score in Estimating Dose-Response Functions." *Biometrika* 87(3): 706-710.

Joffe, Marshall M. and Paul R. Rosenbaum. 1999. "Invited Commentary: Propensity Scores." *American Journal of Epidemiology* 150(4): 327-333.

Kaestner, Robert. 1997. "The Effects of Cocaine and Marijuana Use on Marriage and Marital Stability." *Journal of Family Issues* 18(2): 145-173.

Laumann, Edward O., John H. Gagnon, Robert T. Michael, and Stuart Michaels. 1994. *The Social Organization of Sexuality: Sexual Practices in the United States*. Chicago: the University of Chicago Press.

Lechner, Michael. 1999. "Identification and Estimation of Causal Effects of Multiple Treatments Under the Conditional Independence Assumption." *IZA Discussion Paper* No. 91.

Lechner, Michael. 2002. "Some Practical Issues in the Evaluation of Heterogeneous Labor Market Programs by Matching Methods." *Journal of Royal Statistical Society. Series A* 165(1): 59-82.

Lechner, Michael. 2002. "Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies." *Review of Economics and Statistics* 84(2): 205-220.

Lee, Myoung-Jae. 2005. *Micro-Econometrics for Policy, Program, and Treatment Effects*. New York, NY: Oxford University Press.

Lillard, Lee A., Michael J. Brien, and Linda J. Waite. 1995. "Premarital Cohabitation and Subsequent Marital Dissolution: A Matter of Self-Selection?" *Demography* 32(3): 437-457.

Lillard, Lee A. and Constantijn W.A. Panis. 1996. "Marital Status and Mortality: the Role of Health." *Demography* 33(3): 313-327.

Lillard, Lee A. and Linda J. Waite. 1995. "'Til Death Do Us Part: Marital Disruption and Mortality." *American Journal of Sociology* 100(5): 1131-1156.

Manski, Charles F. 1997. "Monotone Treatment Response." *Econometrica* 65(6): 1311-1334.

Martino, Steven C., Rebecca L. Collins, and Phyllis L. Ellickson. 2004. "Substance Use and Early Marriage." *Journal of Marriage and the Family* 66(2): 244-257.

Orzechowski, Shawna and Peter Sepielli. 2003. *Net Worth and Asset Ownership of Households: 1998 and 2000.* Household Economic Studies. U.S. Census Bureau P70-88.

Power, Chris, Bryan Rodgers, and Steven Hope. 1999. "Heavy Alcohol Consumption and Marital Status: Disentangling the Relationship in a National Study of Young Adults." *Addiction* 94(10): 1477-1487.

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.

Rosen, Sherwin. 1983. "Specialization and Human Capital." *Journal of Labor Economics* 1(1): 43-49.

Teachman, Jay D. 2003. "Premarital Sex, Premarital Cohabitation, and the Risk of Subsequent Marital Dissolution Among Women." *Journal of Marriage and the Family* 65(2): 444-455.

Umberson, Debra. 1987. "Family Status and Health Behaviors: Social Control as a Dimension of Social Integration." *Journal of Health and Social Behavior* 28(3): 306-319.

Umberson, Debra. 1992. "Gender, Marital Status and the Social Control of Health Behavior." *Social Science and Medicine* 34(8): 907-917.

Willis, Robert J. and Robert T. Michael. 1994. "Innovation in Family Formation: Evidence on Cohabitation in the United States." In *The Family, the Market and the State in Ageing Societies*, eds. by John Ermisch and Naohiro Ogawa, pp. 9-45. Oxford, UK: Clarendon Press.

Wilson, Sven E. 2002. "The Health Capital of Families: An Investigation of the Inter-Spousal Correlation in Health Status." *Social Science and Medicine* 55(7): 1157-1172.

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

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Table 3: Any Hazardous Drinking

	MarStat Only	MStat & History	MarStat & DMG	(2) + (3)	School & Work	Educ. Attain.	PriorBeh & Peer	Family Structure	Parental Char.	Parental Control
NevMarr NotCohab	0.156***	0.140***	0.168***	0.169***	0.181***	0.173***	0.156***	0.169***	0.128***	0.182***
NevMarr Cohab	0.120***	0.095***	0.127***	0.098***	0.101***	0.099***	0.099***	0.095***	0.057	0.116***
Disrupted NotCohab	0.153**	0.165***	0.166***	0.157**	0.201***	0.163**	0.139**	0.161**	0.120	0.254**
Disrupted Cohab	0.063	0.050	0.066	0.019	0.036	0.022	0.013	0.005	-0.053	-0.014
Remarried	0.049	0.089	-0.067	-0.067	-0.053	-0.064	-0.113	-0.056	-0.068	(dropped)
Cohab #1		0.076***		0.066***	0.069***	0.069***	0.026	0.074***	0.090***	0.071***
Cohab #2+		0.103***		0.085***	0.098***	0.092***	0.028	0.093***	0.117***	0.055
Early Marr.		-0.009		0.031	0.031	0.034	0.002	0.035	0.023	0.006
Marr. Dur.		0.016***		0.005	0.010	0.005	0.004	0.004	0.005	0.010
BioKid HH		-0.189***		-0.092***	-0.095***	-0.090***	-0.092***	-0.090***	-0.112***	-0.105***
Male			0.112***	0.102**	0.102***	0.104***	0.092**	0.102***	0.101***	0.059***
Age			0.019***	0.018***	0.022***	0.014***	-0.002	0.017***	0.015**	0.031***
Black			-0.245***	-0.229***	-0.202***	-0.229***	-0.192***	-0.218***	-0.187***	-0.242***
Hispanics			-0.024	-0.016	-0.000	-0.014	-0.001	-0.012	0.028	-0.009
NetW50-80			0.065***	0.060***	0.058***	0.058***	0.060***	0.057***	0.043**	0.067***
NW5thQnt			0.097***	0.095***	0.100***	0.092***	0.107***	0.089***	0.076***	0.108***
Enrolled04					-0.022	-0.033				
FullTime04					0.123***					
PartTime04					0.098***					
CuTimeY					0.015***					
CuTime20+					-0.013					
Dropout						-0.008				
GED						-0.029				
SC						0.018				
BA+						0.028				
EvrDrk97							0.110***			
HzDrk97							0.096***			
EvSmok97							0.082***			
PrSmok97							0.012			
PrDrk97							-0.016			
StepFam97								-0.005		
SglMom97								-0.022		
PrContMar									-0.037**	
PrNevMarr									-0.075**	
PrTeenBirth									-0.001	
MomLSHS									-0.046**	
MomSC									-0.001	
MomBA									0.028	
PrNotRelig									0.030	
PrVeryRelg									-0.068**	
MomStrict										0.027
N	7253	7240	5222	5212	4790	5210	5096	5156	3634	3101
Adj-RSq	0.0103	0.0311	0.0921	0.0970	0.1063	0.0969	0.1266	0.0983	0.1046	0.1013

*: p < 0.10; **: p < 0.05; ***: p < 0.10.

Table 5: Descriptive Regressions for Drinking – Final Results

	Current Drk	HazardDrk
Never Married Not Cohabiting	0.095***	0.147***
Never Married Cohabiting	0.080**	0.097***
Disrupted Not Cohabiting	0.211**	0.198**
Disrupted Cohabiting	0.191*	-0.025
Remarried	0.404	-0.088
Cohab once	0.054**	0.042*
Cohab twice +	0.016	0.068*
Early marriage	-0.069	-0.014
Marital duration	0.008	0.014
Biological kids in household	-0.131***	-0.098***
Male	0.041**	0.097***
Age	0.006	0.004
Black	-0.098***	-0.143***
Hispanics	0.029	0.070**
Median-80 net worth	0.042**	0.043**
5th quintile net worth	0.069***	0.089***
Enrolled04	0.001	-0.029
FullTime04	0.106***	0.116***
PartTime04	0.079***	0.090***
Cumulative time in youth	0.013**	0.007
Cumulative time since 20	0.000	-0.016
Dropout	-0.032	0.000
GED	-0.064*	-0.043
SC	0.024	0.004
BA+	0.025	-0.006
Ever drink 97	0.097***	0.111***
Hazard drink 97	-0.003	0.062*
Ever smoke 97	0.054***	0.099***
Step/blended family 97	0.027	-0.006
Single-mother family 97	-0.003	0.002
Parent continuously married	-0.010	-0.049*
Parent never married	-0.048	-0.069*
Mom less than high school	-0.045**	-0.037
Mom some college	0.027	0.003
Mom BA+	0.090***	0.053**
Parent not religious	0.016	0.028
Parent very religious	-0.038**	-0.052***
Mom strict	0.009	0.038**
N	3418	3372
Adj-R-Square	0.1248	0.1364

*: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.10$.

Table 6: Variables Used in 2SLS Estimation
– Based on Sample Correlation Coefficients

	Marriage	Fertility	Drinking
Male, Age, Race/Ethnicity	X	X	X
Geographical regions	X		X
Socioeconomic Status in 1997		X	X
Ever Marijuana 1997	X	X	X
Current Marijuana 1997			X
Antisocial Behavior: Runaway		X	
Antisocial Behavior: Destroy Property			X
Peer_Church 97		X	
Peer_Smoke 97	X	X	
Peer_Drink, Drug 97	X	X	
Peer_SkipSchool, Gang 97		X	X
Peer_NoCollege, NeighborhoodGang 97		X	
Enrollment Status in 2004	X	X	
Educational Attainment in 2004	X	X	
Employment Status in 2004			X
Employment History in 2004	X	X	X
Cumulative weeks worked in teen		X	X
Cumulative weeks worked since age 20	X		X
Youth Health in 1997		X	
Family Structure in 97: Intact		X	X
Parents' Family Structure: Non-Intact		X	
Parents never-married		X	X
Parents gave births in adolescence		X	
Maternal education		X	X
General health of parents		X	X
Parental religiosity			X

Table 7: Variables Used in the Estimation of Propensity Scores (Models 1a, 2a) and Mahalanobis/Propensity Score Matching (Models 1a, 1b, 2a, 2b)

	Model 1a		Model 2a	
	Marr/Cohab	Fertility	Marr/Cohab	Fertility
Male, Age, Race/Ethnicity	X	X	X	X
Geographical regions	X		X	
Socioeconomic Status in 1997		X	X	X
Ever Marijuana 1997	X	X	X	X
Antisocial Behavior: Runaway		X		X
Antisocial Behavior: Destroy Property			X	
Peer_Church 97		X		X
Peer_Smoke 97	X	X	X	X
Peer_Drink, Drug 97	X	X	X	X
Peer_SkipSchool, Gang 97		X		X
Peer_NoCollege, NeighborhoodGang 97		X		X
Enrollment Status in 2004	X	X	X	X
Educational Attainment in 2004	X	X	X	X
Employment Status in 2004				X
Employment History in 2004:				
- Cumulative weeks worked in teen		X	X	X
- Cumulative weeks worked since age 20	X		X	
Youth Health in 1997		X		X
Family Structure in 97: Intact		X	X	X
Parents' Family Structure: Non-Intact		X		
Parents never-married		X		
Parents gave births in adolescence		X		
Maternal education		X	X	X
General health of parents 97		X		X
Parental religiosity 97			X	
Youth religious affiliation 97			X	

Additional controls:	Model 1b	Model 2b
Socioeconomic Status in 1997	X	
Antisocial Behavior: Destroy Property	X	
Cumulative weeks worked in teen	X	
Family Structure in 97: Intact	X	
Ever drink in 97	X	X
Ever binge drink in 97	X	X
Ever smoke in 97	X	X
Maternal education	X	
Parental religiosity 97	X	X
Number of prior cohabiting unions	X	X

Table 8: 2SLS results of hazardous drinking
With propensity scores as regressors, specification 1

```
Linear regression                                      Number of obs =    2839
                                                       F( 44,  2251) =   12.89
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.1412
Number of clusters (HHID) = 2252                       Root MSE      =  .46705
```

HBeh04_Hz~uu	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Mar~r_LPM_s1	**-.116419**	.3291173	**-0.35**	0.724	-.7618241	.5289861
Mar~b_LPM_s1	**.1242436**	.472723	**0.26**	0.793	-.8027749	1.051262
MargPr_Fer~1	**-.178051**	.2615904	**-0.68**	0.496	-.6910347	.3349326
MarHist_Co~1	.0063786	.0250258	0.25	0.799	-.0426975	.0554547
MarHist_Co~2	.0639428	.0447298	1.43	0.153	-.0237732	.1516589
Dmg_Male	**.0896706**	.0311181	**2.88**	0.004	.0286474	.1506938
Dmg_Age04	.0067101	.0239962	0.28	0.780	-.040347	.0537671
Dmg_RacEthnB	-.0771777	.0517616	-1.49	0.136	-.1786831	.0243278
Dmg_RacEthnH	-.0028423	.034657	-0.08	0.935	-.0708054	.0651207
Dmg_RacEthnO	-.1186171	.0608349	-1.95	0.051	-.2379154	.0006812
geo04_Nort~t	.0054805	.0482634	0.11	0.910	-.0891649	.1001258
geo04_South	-.003012	.0301547	-0.10	0.920	-.062146	.0561219
geo04_West	-.0028826	.0313332	-0.09	0.927	-.0643277	.0585624
SES97_HHNe~g	.0053845	.0441621	0.12	0.903	-.0812181	.0919871
SES97_HHN~80	.0471442	.0289121	1.63	0.103	-.0095529	.1038413
SES97_HHNe~t	**.0761491**	.037363	**2.04**	0.042	.0028796	.1494186
HBeh_EvDrk97	**.1208855**	.0254533	**4.75**	0.000	.0709711	.1707999
HBeh_Hz~lu97	.0658983	.0396966	1.66	0.097	-.0119474	.1437441
HBeh_Ever~97	**.0931094**	.0251202	**3.71**	0.000	.0438482	.1423706
RskB_Ever~97	.0141485	.0551313	0.26	0.797	-.093965	.1222619
RskB_Curr~97	.0186904	.050202	0.37	0.710	-.0797567	.1171375
AntiSB_De~97	.0248487	.023593	1.05	0.292	-.0214176	.0711151
Peer_Skip~97	-.0147451	.0166158	-0.89	0.375	-.0473289	.0178387
Peer_Gang97	-.0047507	.0195894	-0.24	0.808	-.0431658	.0336644
EmpSt04_FT~k	**.1553656**	.0448373	**3.47**	0.001	.0674388	.2432924
EmpSt04_PT~k	**.1326986**	.0352143	**3.77**	0.000	.0636427	.2017544
EmpHi~ksTeen	.0001138	.0001536	0.74	0.459	-.0001874	.0004149
EmpHis~ksS20	-.0004361	.0003995	-1.09	0.275	-.0012195	.0003474
FamStr97_i~t	.0386063	.0233711	1.65	0.099	-.0072248	.0844375
Fam_ParNev~r	.0511478	.0378859	1.35	0.177	-.0231472	.1254428
Par~oMomLSHS	-.0519797	.0368538	-1.41	0.159	-.1242506	.0202912
ParEd_BioM~C	-.0087964	.0257573	-0.34	0.733	-.059307	.0417142
ParEd_BioM~A	.0417752	.0298402	1.40	0.162	-.0167421	.1002925
HEAP_GenHe~G	-.0147636	.0237401	-0.62	0.534	-.0613182	.0317911
HEAP_GenHe~r	-.0178785	.039003	-0.46	0.647	-.0943641	.0586072
ParRelg_No~g	**.0636726**	.0265428	**2.40**	0.017	.0116216	.1157236
ParRelg_Ve~g	-.0275092	.0237361	-1.16	0.247	-.0740562	.0190378
YRLG97_Pro~m	**-.1324976**	.0305892	**-4.33**	0.000	-.1924837	-.0725116
YRLG97_Prot3	-.0747616	.0537095	-1.39	0.164	-.180087	.0305638
YRLG97_Cath	-.004439	.0294477	-0.15	0.880	-.0621865	.0533085
YRLG97_Athe	-.0083241	.116578	-0.07	0.943	-.2369357	.2202875
YRLG97_Per~l	-.0192835	.0757629	-0.25	0.799	-.1678559	.1292889
YRLG97_NoRLG	-.129501	.0723878	-1.79	0.074	-.2714547	.0124527
YRLG97_Oth~s	-.1823362	.0914326	-1.99	0.046	-.3616373	-.0030352
_cons	.1740907	.4939085	0.35	0.725	-.794473	1.142654

Table 9: Summary descriptives of marginal probabilities (propensity scores)

Model 1	N	Median	Mean	S. D.	Min	Max
Marginal Prob - Cohab	6,893	0.146	0.161	0.091	0.012	0.554
Marginal Prob - Marriage	6,893	0.104	0.133	0.109	0.000	0.640
Marginal Prob - Nev Mar.	6,893	0.732	0.706	0.171	0.182	0.985
Marginal Prob - Fertility	4,594	0.148	0.219	0.211	0.001	0.937
Model 2						
Marginal Prob - Cohab	3,581	0.135	0.157	0.106	0.005	0.688
Marginal Prob - Marriage	3,581	0.078	0.118	0.120	0.000	0.814
Marginal Prob - Nev Mar.	3,581	0.760	0.725	0.179	0.128	0.994
Marginal Prob - Fertility	4,748	0.148	0.220	0.211	0.000	0.939

Table 10: Mahalanobis matching results

	Marriage					Cohabitation				
Sample	Treated	Controls	Diff.	S. E.	T-Stat	Treated	Controls	Diff.	S. E.	T-Stat
Model 1a										
Unmatched	0.344	0.478	-0.134	0.022	-5.97	0.451	0.461	-0.010	0.021	-0.48
ATT	0.344	0.476	**-0.132**	0.032	**-4.19**	0.451	0.432	**0.019**	0.029	**0.66**
ATU	0.478	0.323	-0.155			0.461	0.457	-0.004		
ATE			**-0.152**					**-0.000**		
N	567	3,625				676	3,516			
Model 1b										
Unmatched	0.330	0.469	-0.139	0.027	-5.19	0.441	0.455	-0.013	0.024	-0.55
ATT	0.330	0.438	**-0.108**	0.037	**-2.95**	0.441	0.502	**-0.061**	0.036	**-1.68**
ATU	0.469	0.265	-0.204			0.455	0.452	-0.003		
ATE			**-0.192**					**-0.012**		
N	388	2,860				512	2,736			
Model 2a										
Unmatched	0.327	0.470	-0.143	0.026	-5.40	0.439	0.455	-0.016	0.024	-0.68
ATT	0.327	0.417	**-0.090**	0.038	**-2.36**	0.439	0.433	**0.006**	0.033	**0.17**
ATU	0.470	0.328	-0.141			0.455	0.460	0.005		
ATE			**-0.135**					**0.005**		
N	398	2,949				526	2,821			
Model 2b										
Unmatched	0.327	0.469	-0.143	0.026	-5.40	0.439	0.455	-0.016	0.024	-0.67
ATT	0.327	0.425	**-0.098**	0.037	**-2.63**	0.439	0.462	**-0.023**	0.0402	**-0.57**
ATU	0.469	0.390	-0.080			0.455	0.425	-0.030		
ATE			**-0.082**					**-0.029**		
N	398	2,948				526	2,820			

Table 11: Pseudo-R2 and LR-Chi2 from Mahalanobis matching

	Marriage			Cohabitation		
	Pseudo-R2	LR Chi2	p > Chi2	Pseudo-R2	LR Chi2	p > Chi2
Model 1a						
Unmatched	0.135	447.03	0.000	0.071	264.62	0.000
Matched	0.139	463.35	0.000	0.080	296.51	0.000
Model 1b						
Unmatched	0.166	393.58	0.000	0.408	1155.75	0.000
Matched	0.059	139.64	0.000	0.085	239.25	0.000
Model 2a						
Unmatched	0.164	401.30	0.000	0.094	274.53	0.000
Matched	0.186	453.13	0.000	0.112	325.88	0.000
Model 2b						
Unmatched	0.173	421.48	0.000	0.408	1186.96	0.000
Matched	0.136	332.29	0.000	0.127	369.80	0.000

Table 12: Median and mean absolute biases before & after Mahalanobis matching

	Marriage			Cohabitation		
	Median	Mean	S.D.	Median	Mean	S.D.
Model 1a						
Before	9.579	15.080	13.505	12.160	14.668	9.139
After	4.136	4.280	3.437	3.262	4.064	2.994
Δ in std. bias	-0.494	-0.129		0.241	0.248	
Model 1b						
Before	11.839	16.469	14.328	15.486	21.439	30.925
After	5.081	7.480	7.511	4.627	7.184	10.984
Δ in std. bias	0.150	0.154		0.080	0.039	
Model 2a						
Before	10.292	15.406	13.833	15.268	15.972	10.933
After	4.824	5.104	3.384	4.366	4.735	3.387
Δ in std. bias	-0.682	-0.395		0.107	0.063	
Model 2b						
Before	11.036	15.772	14.041	15.937	20.449	30.284
After	3.610	4.680	4.656	4.077	4.833	4.582
Δ in std. bias	0.011	0.118		-0.364	-0.380	

Before and after: absolute bias obtained by pstest command.
Δ in std. bias = change in standardized bias after matching.
Δ in std. bias = (Std. bias before) - (Std. bias after).

Table 13: Variables with biases increase by 20% after Mahalanobis matching

	Marriage	Cohabitation		Marriage	Cohabitation
Model 1a	Race/ethnicity - O.	Race/ethnicity - O.	Model 2a	Race/ethnicity - O.	
		Race/ethnicity - B.			Race/ethnicity - B.
		Geogr04_West			
	Youth health '97	Youth health '97		Youth health '97	
	Parent health '97			Parent health '97	
		MomEd_SomeColl		MomEd_SomeColl	MomEd_SomeColl
				SES97	
				EduSta04_Dropout	
				YRLG97_Prot3	
				YRLG97_Others	
					YRLG97_Catholic
Model 1b	Race/ethnicity - O.	Race/ethnicity - O.	Model 2b	Race/ethnicity - O.	Race/ethnicity - O.
					Race/ethnicity - H.
		Geogr04_NorthE.			Geogr04_West
	Empl Hist - Teen				
	SES97			SES97	SES97
	Youth health '97	Youth health '97		Youth health '97	Youth health '97
	Parent health '97			Parent health '97	
				MomEd_SomeColl	MomEd_SomeColl
				EduSta04_Dropout	
					EduSta04_GED
	Parent Non-intact				
	Parent Never-Marr				
	Parent Teen-Birth				
		MomEd_BA			
				YRLG97_Prot3	
				YRLG97_Others	
					YRLG97_Catholic

Table 14: Pair-wise comparisons using two definitions of propensity score

	Marriage vs. Never-Married			Cohabitation vs. Never-Married		
	Full	MN(1)	MN(2)	Full	CN(1)	CN(2)
Model 1						
ATT	**-0.153**	**-0.168**	**-0.178**	-0.029	-0.038	-0.048
S.D.	0.029	0.031	0.031	0.027	0.028	0.028
T-stat	**-5.236**	**-5.412**	**-5.744**	-1.067	-1.346	-1.700
ATE	-0.193	-0.203	-0.206	-0.005	-0.024	-0.021
Treated N - On Support	567	567	565	673	673	674
Total N - On Support	4,035	3,383	3,390	4,092	3,535	3,536
Total N	4,192	3,516	3,516	4,192	3,625	3,625
% On Support	96.25%	96.22%	96.42%	97.61%	97.52%	97.54%
Model 2						
ATT	**-0.168**	**-0.170**	**-0.156**	-0.057	**-0.085**	-0.020
S.D.	0.037	0.039	0.039	0.032	0.033	0.033
T-stat	**-4.580**	**-4.370**	**-4.031**	-1.821	**-2.596**	-0.607
ATE	-0.141	-0.154	-0.155	-0.005	-0.027	-0.037
Treated N - On Support	388	383	384	521	520	524
Total N - On Support	3,071	2,578	2,542	3,071	2,895	2,899
Total N	3,347	2,821	2,821	3,347	2,949	2,949
% On Support	91.75%	91.39%	90.11%	91.75%	98.17%	98.30%

Full: full sample.
MN: Sample consisted only of the married and the never-married respondents.
CN: Sample consisted only of the cohabitors and the never-married respondents.
(1): Propensity score = marginal probability (marriage or cohabitation) from multinomial probit.
(2): Propensity score = MP(marriage or cohab)/[MP(marriage or cohab) + MP(never-married)],
MP = marginal probability.

Table 15: Cell sizes in three 3x3 matrix

First tercile of marginal probability of being a parent (least likely)

P2 = 3	11	43	20
P2 = 2	84	226	133
P2 = 1	683	276	58
N	P1 = 1	P1 = 2	P1 = 3

P1: Marginal probability for marriage.
P1 = 1: lowest third, least likely to marry; P1 = 3, highest third, most likely to marry.
P2: Marginal probability for cohabitation.
P2 = 1: lowest third, least likely to cohabit; P2 = 3: highest third, most likely to cohabit.

Second tercile of marginal probability of being a parent

P2 = 3	56	172	189
P2 = 2	173	280	241
P2 = 1	271	95	57
N	P1 = 1	P1 = 2	P1 = 3

P1: Marginal probability for marriage.
P1 = 1: lowest third, least likely to marry; P1 = 3, highest third, most likely to marry.
P2: Marginal probability for cohabitation.
P2 = 1: lowest third, least likely to cohabit; P2 = 3: highest third, most likely to cohabit.

Third tercile of marginal probability of being a parent (most likely)

P2 = 3	92	315	672
P2 = 2	108	137	159
P2 = 1	18	16	12
N	P1 = 1	P1 = 2	P1 = 3

P1: Marginal probability for marriage.
P1 = 1: lowest third, least likely to marry; P1 = 3, highest third, most likely to marry.
P2: Marginal probability for cohabitation.
P2 = 1: lowest third, least likely to cohabit; P2 = 3: highest third, most likely to cohabit.

Table 16: Subclassification/stratification results

	Marriage	Cohabitation
ATT	**-0.174**	**-0.016**
S.D.	0.032	0.028
T-Stat	**-5.512**	**-0.580**
ATE	**-0.179**	**0.007**
Total-N On Support	3,715	3,957
Total-N	4,597	4,597
% On Support	80.81%	86.08%
Number of cells unused	3	3
N - cases in these cells	114	39