The Effect of Marital Status on Home Ownership among Low-Income Households

Michal Grinstein-Weiss  
*University of North Carolina at Chapel Hill*

Pajarita Charles  
*University of North Carolina at Chapel Hill*

Shenyang Guo  
*University of North Carolina at Chapel Hill*

Kim Manturuk  
*University of North Carolina at Chapel Hill*

Clinton Key  
*University of North Carolina at Chapel Hill*

This research examines whether married low-income renters are more likely to become home owners than comparable single, low-income renters. To do so, it employs data from the Community Advantage Panel Study and discrete-time survival analysis with propensity-score matching. Results suggest that married couples buy homes at higher rates, and buy them more quickly, than do their unmarried counterparts. Estimates in models that use propensity-score matching are robust to the control of selection bias between the married and the unmarried groups. The findings suggest that efforts to encourage marriage among low-income couples may be associated with subsequent economic mobility through home ownership.

The promotion of home ownership is one strategy that policy makers use to improve the well-being of disadvantaged families. A home is a valuable asset, and for multiple reasons, home ownership is an important area of study. Home ownership remains a long-term goal for many Amer-
icans, in spite of the recent downturn. However, the road to home ownership is more difficult than ever; it is harder to get credit for the purchase of a home, and underwriting standards have become stricter. Given scarce resources and current credit constraints, low-income married couples may have an easier time than comparable single people in achieving and sustaining home ownership. Understanding the role of marital status in the transition to home ownership will help policy makers to target policy that pursues the interests of all people.

There also is evidence that home ownership confers considerable social benefits to low- and moderate-income householders if it is done right. Ownership may increase community involvement and psychological well-being (Manturuk, Lindblad, and Quercia 2009; Manturuk and Quercia 2010; Ding et al. 2011). Understanding the relationship between marriage and home ownership may suggest ways to deepen these beneficial outcomes and reduce risk of divorce for vulnerable couples.

Moreover, research suggests that home ownership provides a wealth-building opportunity for low-income families if borrowers' financial needs are properly aligned with the interests of mortgage lenders (Abromowitz and Ratcliffe 2010). Therefore, understanding the factors that influence tenure choice (i.e., the decision between owning and renting one’s home) has important implications for policy makers and researchers seeking to improve the financial welfare of families. One such factor is marriage.

Research rarely examines marriage as a key predictor of home ownership, and efforts to construct tenure choice models find that marriage plays a less influential role than economic factors. Early research tended to focus on tax incentives and permanent income as the main predictors of home ownership (King 1980; Haurin 1991; Poterba 1992). Recent explanations of tenure choice decisions emphasize that barriers to borrowing impede home ownership (Haurin, Hendershott, and Wachter 1997; Rosenthal 2001; Calem, Firestone, and Wachter 2010). Income, wealth, and asset holdings often limit a low-income buyer’s access to credit, as such financial conditions may impede one’s ability to meet deposit requirements or to make mortgage payments (Hendershott et al. 2009).

There are numerous financial and nonfinancial reasons why one would expect marital status to be relevant in tenure choices. Life-course “triggers,” such as marriage and childbirth, are found to be strong predictors of movement within the housing market (Clark and Huang 2003, 323). Furthermore, the attributes of married couples could lend themselves to stronger preferences for home ownership than do the attributes of unmarried couples. For example, married couples may be less inclined to move frequently because both partners may have ties to the area, and they would therefore have a stronger preference for home ownership than a one-person household. Also, households that shift
from renting to home ownership tend to consist of families or couples whose relationship has reached a certain level of stability (Clark, Deurloo, and Dieleman 1994). This stability is likely to be derived in part from marriage and from the financial capability that often accompanies marriage (Hendershott et al. 2009). Expanding understanding of the effect of family structure on tenure choice decisions is therefore important because it may inform future housing policies, particularly policies for low- and moderate-income households that suffered home ownership setbacks in the recent housing crisis.

Background and Conceptual Framework

Tenure Choice Models

Models of tenure choice focus on factors that affect consumers’ choice between owning a home and renting one. Some tenure choice research also focuses on a variant of these models, examining the role of home ownership subsidies in marital outcomes (Eriksen 2010). A well-established topic in the economic literature, tenure choice is characteristically examined as a function of rational or economic determinants (Hendershott et al. 2009).

In rational economic terms, the choice of home ownership over renting is merely a cost-saving measure. People elect to own a home when the overall cost of doing so is less than the net cost of renting. Models of tenure choice include numerous financial determinants of owner-occupied housing, including factors related to tax incentives. The deduction of mortgage interest and property tax from income for federal tax purposes can, in effect, reduce the cost of owning compared to renting (Rosen 1979; Poterba 1992). Permanent income, which is derived from the capacity to accrue long-term earnings, is also considered important in tenure choice because owning a home is usually tied to making future mortgage payments. As such, prospective buyers consider their potential for future income when evaluating tenure choice options (King 1980). Sometimes modeled using proxies such as education, labor market history, and future wage rates, permanent income is found to be positively associated with home ownership (Linneman and Wachter 1989; Haurin 1991; Haurin et al. 1997).

Recent literature on tenure choice focuses on borrowing constraints imposed on prospective buyers by individual limitations, such as limitations in wealth or current income, or by systemic issues like credit barriers and the absence of affordable homes (Jones 1995; Goodman and Nichols 1997; Haurin et al. 1997; Rosenthal 2001; Calem et al. 2010). The primary literature on borrowing constraints and home ownership consistently finds that insufficient income and wealth are strongly and negatively associated with rates of home ownership (Haurin et al. 1997).
The literature also examines lack of a credit history and impaired credit as potential borrowing constraints that negatively affect one’s tenure options (Calem et al. 2010). In general, research on tenure choice demonstrates that the presence of borrowing constraints reduces the probability of choosing to own a home (Jones 1995; Goodman and Nichols 1997; Haurin et al. 1997; Rosenthal 2001; Calem et al. 2010).

Home Ownership among Low-Income Households

Research suggests that home ownership is associated with benefits for individuals and communities. These benefits include savings, wealth (Skinner 1989; Di, Yang, and Liu 2003), social and civic involvement (Dreier 1994; DiPasquale and Glaeser 1999; Manturuk et al. 2009), and positive child outcomes (Boehm and Schlottmann 1999; Haurin, Parcel, and Haurin 2002; Holupka and Newman 2008). Despite these potential advantages, home ownership is unattainable for many (Withers 1998), and it is especially elusive for those who are economically disadvantaged. To address this gap, policy makers promote home ownership through targeted programs for low-income and minority households (Scanlon 1998). In the 1990s, these programs helped to drive an increase in the rate of home ownership among low-income individuals (Belsky and Duda 2002).

Despite the success of some of these programs, home ownership among low-income individuals has become a questionable strategy for achieving economic well-being, and doubts are raised about the social benefits of home ownership among low-income households (Engelhardt et al. 2010). Criticism of efforts to promote home ownership has become more pronounced with the dramatic increase in foreclosures during the recent housing crisis.

However, evidence suggests that low-income families benefit from well-designed home ownership programs, many of which have proven successful despite the foreclosure crisis. A recent study of foreclosure rates among affordable-housing purchasers in five urban areas finds that program participants had lower-than-average rates of mortgage default across all five cities (Reid 2009). In fact, low-wealth home buyers have no greater risk of default than their higher-wealth counterparts, provided the low-wealth purchasers hold prime rate mortgages with conventional terms (Ding et al. 2011). In the aftermath of the downturn, however, it is insufficient to focus evaluation efforts solely on whether policies to promote low-income home ownership are not detrimental.

It is also important to consider whether home ownership continues to provide purchasers with a way to build wealth and assets. It will be difficult to answer this question fully until the recovery is further along, but preliminary evidence suggests that responsible home ownership does build wealth, even during a market downturn (Riley and Quercia
Examination of the factors that affect tenure choice therefore continues to be a worthwhile undertaking.

Marriage and Tenure Choice

Two dominant perspectives are often used to explain the role that marriage plays in the decision to purchase a home. Sociological theory suggests that several attributes may make home ownership more likely among married couples than among their unmarried counterparts. First, married couples may be less likely to feel that their lives are regularly in transition than singles and, hence, more willing to make lasting, long-term decisions. Married couples are more likely than singles to be financially stable and are less likely to anticipate moving again in the near future. As such, married couples tend to be more willing to commit to the idea of purchasing a home (Clark et al. 1994; Mulder and Wagner 2001). Second, expectations inextricably tied to married life influence such important decisions as how many children to have and whether to purchase a home (Lupton and Smith 2003). People generally have strong preferences for home ownership but prefer to wait until they achieve stability in their income and their family situation (Clark et al. 1994). This stability often comes in the form of marriage and is linked to levels of commitment among partners (Smits and Mulder 2008); households with more highly committed members are more likely to purchase a home in anticipation of future events, such as having children (Feijten and Mulder 2002). Third, people often make the shift from renting to home ownership after experiencing an important life transition (Clark et al. 1994; Deurloo, Clark, and Dieleman 1994). Marriage is one life event that could have this type of triggering effect.

An economic perspective suggests financial reasons why married couples may be more apt to buy a home than are single people. First, married couples tend to have greater financial capability (i.e., higher combined assets and income) than singles. This capability makes them more likely than singles to purchase a home (Plaut 1987; Linneman and Wachter 1989; Hendershott et al. 2009). Second, married couples are likely able to save more than single individuals and are likely to have higher levels of net worth because couples can benefit from economies of scale by sharing expenses and purchasing goods and services in efficient sizes; couples also can benefit from labor market specialization (Grinstein-Weiss, Zhan, and Sherraden 2006). Third, married men generally earn more than unmarried men (Waite and Gallagher 2000). Fourth, marriage can be positively associated with social capital, the accumulation of which can result in opportunities that lead to saving. Finally, married couples commonly have access to benefits, like health and life insurance, that promote savings and enable them to consider home ownership opportunities (Lupton and Smith 2003; Grinstein-
These economic considerations lead us to expect that marriage can help potential home owners overcome borrowing constraints that limit ownership opportunities. We therefore expect that marriage will be found to increase the rate of home ownership.

The limited tenure choice literature suggests that marriage can indeed play a role in predicting home ownership, as can income, education, race, number of children, and age (Mulder and Wagner 1998; Andrew, Haurin, and Munasib 2006; Hendershott et al. 2009). However, these studies are flawed in several ways. First, they do not successfully address the problem of bias in estimates that result from selecting into marriage. To draw causal inference about the role of marriage in tenure choice decisions, methods must be more rigorous than the conventional covariance control approach. This article addresses the issue of selection effects by using propensity-score matching to examine the potentially causal role that marriage plays in the decision to buy a home. A second critique of previous studies is that they use data drawn from countries outside the United States. Tenure choice decisions may operate differently in other countries than they do in the United States because of dissimilarities in economic incentives for home ownership. For example, tax deductions are offered to home owners in the United States, but other countries (e.g., Australia) provide cash subsidies for mortgage payments and down payments (Bourassa and Yin 2006). Tax laws in other countries may even favor renting over home ownership or may bias households against owning altogether (Henderson and Ioannides 1983; Jones 1995).

A third critique is that few studies focus on tenure choice decisions among low-income populations. Because of this, the nature of low-income households’ relationship to home ownership remains uncertain. This uncertainty is problematic if one considers the role that marriage may play in home ownership; some research suggests that marriage in low-income groups manifests itself differently than it does in higher-income groups. These differences imply an unconventional ordering of typical life-course events. For example, Kathryn Edin and Maria Kefalas (2005) discover that marriage norms and conventions in low-income communities are different from those in higher-income communities and that marriage is viewed as an end goal; it is something that comes after one obtains a home, finishes school, finds a job, and bears children. If this is true, then home ownership among single-person, low-income households might come before marriage, not after, as it often does among middle- and high-income U.S. households.

Taken as a whole, the literature on marital status and low-income home ownership is fragmented. No single study analyzes a low-income, U.S.-based sample and employs a rigorous analytical approach to examine explicitly the effect of marriage on tenure change. The purpose of the current research is to address this gap and to improve under-
standing of the role that marital status plays in the decision to become a home owner. Addressing this gap is in part accomplished with the use of data from a U.S.-based longitudinal study of low- and moderate-income individuals who own or rent their homes. Previous research shows that this sample of owners and renters is similar to other nationally representative random samples of low-income groups (Riley and Ru 2009). We hypothesize that married couples will shift from renting to home ownership at higher rates, and at faster rates, than those of their unmarried counterparts.

Method

Data

This study uses data collected as part of an evaluation of the Community Advantage Program, a secondary mortgage-market program developed to underwrite 30-year fixed-rate mortgages for families that would otherwise likely receive a subprime mortgage or be unable to purchase a home. In order to qualify for the program, participants were required to meet one of the following criteria: (1) have an annual income of no more than 80 percent of the area median income (AMI), (2) be a member of a racial or ethnic minority and have an income not in excess of 115 percent of AMI, or (3) purchase a home in a high-minority (greater than 30 percent) or low-income (less than 80 percent of AMI) census tract and have an income not in excess of 115 percent of AMI. By the end of 2004, 28,573 mortgages were funded through the Community Advantage Program.

In 2004, the Community Advantage Panel Study (CAPS) was initiated to assess the effect of home ownership on the lives of Community Advantage Program participants. In order to facilitate this analysis, researchers chose a random sample of the program’s borrowers to participate in annual surveys. A comparison group of renters was matched to this sample of home owners; the groups are matched on neighborhood and income characteristics. Matching is limited to the 30 U.S. metropolitan areas that have the highest number of Community Advantage Program owners. The renter sample was randomly selected from public telephone directory lists of residents who lived within the same census blocks as already-enrolled home owners. Like the Community Advantage Program home owners, the renters had to meet one of the three program eligibility criteria. In addition, participants in the renter sample had to be between 18 and 65 years old and pay rent to the owner of their residence. The final year 1 sample is composed of 3,743 home owners and 1,530 renters. Because this study examines the transition into home ownership, the analysis includes only the 1,530 renters.

Two unique attributes make CAPS data ideally suited for this research.
First, CAPS focuses exclusively on the low- and moderate-income population. Recent policies aimed at promoting home ownership as an asset-building strategy are largely directed at low-income families. Likewise, federal programs to promote and sustain marriage identify reducing poverty as a goal. Therefore, it is important to focus on this population in examining the relationship between marriage and home ownership from a policy perspective. Second, because CAPS participants are interviewed annually, the data set includes frequent data points for time-varying measures. Most other panel data sets, such as the Panel Study of Income Dynamics and the National Longitudinal Study of Youth, interview participants less frequently. Using data collected annually allows us to model the temporal relationship between marriage and home ownership more accurately than would be possible with other data sets.

It is important to note, however, that CAPS does not provide nationally representative samples of all home owners and renters. Rather, CAPS is representative only of low- and moderate-income people living in certain urban areas. In order to assess how the CAPS sample compares to the relevant random national sample, Sarah Riley and Hong Ru (2009) compare the 2004 CAPS sample of home owners with a sample of low-income home owners who participated in the U.S. Census Bureau’s 2004 Current Population Survey. They find that the sociodemographic composition of CAPS participants is very similar to that of the members of the Current Population Survey sample. The one notable difference between respondents to the two surveys is that over 90 percent of CAPS home owners are employed, but this is true of only 70 percent of low-income home owners in the Census Bureau’s sample. This may be because all CAPS owners purchased their homes around the time of the interview, and the loan for the purchase required them to have a steady source of income at that time.

The current analysis uses five waves of data from the CAPS renter panel. Data were collected annually from 2004 to 2008. As with any longitudinal survey, CAPS experienced sample attrition. As of 2008, the CAPS renter panel consisted of 923 (60 percent) of the original 1,530 respondents. Of those 923 cases, complete data are available for 642 respondents on marital status and such important predictors of home ownership as income, race, and age. Applying the multiple imputation approach (Schafer 1997; Little and Rubin 2002), this study imputes missing data for the 281 subjects who had values on one or more independent and matching variables. At the baseline interview in 2004, the distribution of these 923 cases on reported marital status is as follows: 24.5 percent were married, 35.4 percent were never married, 8.8 percent were cohabiting, 26.6 percent were divorced, and 4.7 percent were widowed. The current study excludes people who were cohabiting, divorced, or widowed from the analysis. These exclusions are made for the sake of parsimony and in order to better isolate the effect of marriage on
home ownership from the effect of other forms of romantic involvement on home ownership. It employs a final analytic sample of 553 (i.e., 179 were married at the baseline, and 374 were never married). Using the five implicants (or versions of the data) generated through multiple imputation with chained equations, the study achieves a relative efficiency of 90 percent.

Measures
The key outcome measure in this study is the time that elapses between the first wave of the CAPS survey and the home purchase. Because each respondent began the study as a renter, the outcome variable is measured in discrete units (1–4) as the number of data waves that elapse between the respondent’s entry into the program and the point at which he or she purchased a house.

The key independent variable of this study is marital status. In the Community Advantage Program, marital status is captured with six categories: living with a partner, married, widowed, divorced, separated, or never married. Respondents who indicated that they were with partners, widowed, divorced, or separated are excluded from the analysis so that the study respondents are classified as either married or never married (i.e., truly unmarried singles). Marital status is operationalized as a time-varying covariate; that is, marital status refers to the respondent’s reported status in the survey wave before the time he or she purchased a home (e.g., if it took a respondent 1 year to purchase a home, marital status refers to the reported status at the end of year 1; if it took 2 years, marital status refers to the reported status at the end of year 2). As such, the study investigates the possibility that the relationship between marriage and home ownership is causal, and that investigation meets the requirement of temporal order; namely, the cause should precede the effect.

All models include sociodemographic control variables. Characteristics used in the propensity-score matching are measured at year 1 of participation in the Community Advantage Program. In models addressing time elapsed before home purchase, employment, number of children, and income are used at the value collected nearest to the event, while other covariates use baseline values. Respondents self-report gender (1 = female; 0 = male) and race. Four categories are used to capture race and ethnicity. The analysis reduces these categories to produce two indicator variables (i.e., African American; Hispanic and other) and a reference category (white). Age at baseline is measured in years and recorded as an integer. Data on the level of education completed by the respondent are collected at the baseline interview. This variable is treated as an ordinal variable in the analysis because small cell counts for some tiers make indicator variables problematic in
models. Possible responses on the education measure range from “11th grade or less” (coded as 1) to “graduate degree” (coded as 8). Four respondents indicated their education was nontraditional. The analyses recode their responses to treat those cases as if the education data are missing. The measure of the number of children in the household is constructed using the respondent’s baseline-reported household roster, which includes a count of all child household members (ages 0–17 years). Employment is recoded into an indicator variable (1 = working; 0 = not working) from four response categories: working, not working, out of labor force, and retired. Income is measured in thousands of nominal dollars.

The analysis controls for the characteristics of the respondent’s census tract at baseline. These characteristics include the median house value in the respondent’s census tract, the tract’s median rent, and the tract’s disadvantage score. The tract neighborhood disadvantage score is constructed from several other tract-level items in the 2000 census: the percentage of tract residents who are not working, the percentage in poverty, the percentage on public assistance, and the percentage of tract households headed by single adults with children (Sampson, Raudenbush, and Earls 1997). Research establishes that neighborhood characteristics are associated with the individual decision to purchase a home (Percy, Hawkins, and Maier 1995). This study includes tract-level characteristics to enable the analyses to isolate the effect of marriage on home ownership from the effect of neighborhood context on home ownership. Although most of the variables discussed above are fixed characteristics, income and employment change over time. For the logistic models used to create the propensity scores, all characteristics are measured at baseline. For the subsequent survival analysis, the study uses time-varying measures of employment, number of children, and income.2

Missing data are imputed using multiple imputation through chained equations (Rubin 1987; Van Buuren, Boshuizen, and Knook 1999; Roy- ston 2005), a procedure implemented in the Imputation by Chained Equations package for Stata. Imputation helps to reduce bias caused by nonresponse to survey questions and to reduce such bias more effectively than does the common practice of listwise deletion (Raghunathan 2004). Using multiple imputation rather than single imputation allows researchers to include variability from the imputation. The imputation model includes all variables in the analytic model. For each variable on which information is missing, information is modeled using all of the other variables in the data set. Values for missing data are selected from the resultant conditional density and applied in a new iteration of the data set. Because of the uncertainty inherent in the selection of a value, this process is repeated iteratively. Special attention was given to the imputation model to ensure that the correct functional form is specified
for imputed variables, and postimputation diagnostics compare distributions of imputed values to the distribution of known values. In light of recommendations by Paul Von Hippel (2007), imputed values of the dependent variable are not used in the analysis. Given the fraction of missing data, five imputed data sets were created (Schafer 1997).

**Analysis**

This study aims to test a potentially causal relationship: if the models account for important covariates available to the study, does marital status have a net and strong correlation with the transition into home ownership? In the past 30 years, researchers have recognized the need to develop more efficient approaches for assessing treatment effects in studies based on observational data. This growing interest in the search for consistent and efficient estimators of causal effects led to a surge in efforts to estimate average treatment effects under various sets of assumptions (e.g., Heckman 1978, 1979; Rosenbaum and Rubin 1983).

Researchers find that the conventional covariance control approach has numerous flaws and should be replaced by more rigorous methods for drawing causal inference. For instance, Michael Sobel (1996) criticizes the practice, common in sociology, that uses a dummy variable to evaluate the treatment effect in a regression model (or a regression-type model) of survey data. This article uses the terms “treatment” and “nontreatment” or “control” in a broad sense; that is, they are used under the setting of observational studies and refer to conditions associated with the central cause under study. Specifically, the term “treatment” in this study denotes being married, and the nontreatment (or control) group denotes that the respondent is reported as not married.

The literature indicates that a covariance control approach raises several analytical issues, including exogeneity of the treatment variable, the strongly ignorable treatment assignment assumption, and the difficulties of covariance control in addressing nonignorable treatment assignment. The primary problems discussed in the literature may be summarized as follows. The first of these is found in research suggesting that these models specify the dummy treatment variable as exogenous, but in fact the variable is not. Determinants of incidental truncation or sample selection should be explicitly modeled first, and selection effects should be taken into consideration in estimates of causal effects on outcomes (Heckman 1978, 1979). The second problem is that the strongly ignorable treatment assignment assumption (i.e., conditional on covariates, the treatment assignment is independent from outcomes under both treatment and control conditions) is prone to violation in observational studies. Under such conditions, the presence of the endogeneity problem leads to a biased and inconsistent estimation of the regression coefficient (Rosenbaum and Rubin 1983; Berk 2004; Imbens
A third problem is that covariance control does not automatically correct for nonignorable treatment assignment (Guo and Fraser 2010). To draw valid causal inference, this study applies the Neyman-Rubin counterfactual framework (Rubin 1974, 1990, 2006; Splawa-Neyman 1990; Morgan and Winship 2007), which serves as a conceptual model that guides the data analysis. Under this setting, a counterfactual is a potential outcome or event that would have happened in the absence of the cause (Shadish, Cook, and Campbell 2002). A counterfactual framework emphasizes that individuals selected into either the treatment or the nontreatment group have potential outcomes in both states, that is, the state in which they are observed and the one in which they are not observed. The Neyman-Rubin framework offers a practical way to evaluate counterfactuals. In analysis of data from a sample that represents the population of interest, the standard estimator for the average treatment effect is seen as the difference between two estimated medians from the sample data; that is,

\[ \hat{\tau} = \text{median}(\hat{T}_1|w = 1) - \text{median}(\hat{T}_0|w = 0), \]

where \(\hat{T}_1\) is the survival time under the treated condition, \(\hat{T}_0\) is the survival time under the control condition, and \(w\) is a binary variable indicating treatment receipt (i.e., \(w = 1\), treatment condition; and \(w = 0\), control condition).

This study uses several methods, including a discrete-time survival analysis, a propensity-score optimal-pair matching, and a propensity-score optimal full matching, to balance data, to examine net association, and to draw a valid inference concerning causality. Specifically, the study uses (a) a discrete-time survival analysis applied to the original sample without matching, (b) a propensity-score optimal-pair matching that uses generalized boosted modeling (GBM) to estimate the propensity score and a follow-up discrete-time survival analysis, and (c) a propensity-score optimal full matching that uses GBM to estimate the propensity score and a follow-up Hodges-Lehmann aligned rank test (Hodges and Lehmann 1962), and (d) all analyses are conducted on five files of multiply imputed missing data of independent and matching variables. Results from these data sets are then aggregated using Donald Rubin’s rule for multivariate models (Schafer 1997; Allison 2002; Little and Rubin 2002; Graham 2009).

The authors performed two sensitivity analyses to warrant that the final findings are robust, that they hold in a differently defined population, and that they persist in estimates from a different statistical procedure. The first such analysis aims to check whether the same estimated effect of marriage on home ownership holds in a population that defines the nonmarried category more broadly; to do so, the study replicates all analyses in the sample of 923 cases but treats those who lived with partners, divorced, or are widowed as the same as truly unmarried singles.
(i.e., the estimates contrast 23.2 percent of married cases with 76.8 percent nonmarried cases). The second test aims to check sensitivity of the final results to a more restrictive requirement for a common-support region in matching; to do so, the study applies the greedy propensity-score matching (i.e., the nearest neighbor within caliper matching) to the same data analyzed by the optimal matching.

Findings

After missing-data imputation, the analytic sample contains 553 participants. Of those, 179 (32.4 percent) report that they were married; that is, they reported at the baseline interview that they were married, or they reported being married during the 4-year study period. The remaining 374 participants (67.6 percent) report that they never married at any point during the study period.

Table 1 presents descriptive statistics for the sample as well as the results of imbalance checks conducted before and after matching. As it shows, the overall sample before matching is not balanced on various covariates. For example, the table shows that the original sample contains more married males than married females, and the difference is statistically significant ($p < .01$). Other statistically significant covariates predicting differences on marital status include race and income. If these differences in marital status were not considered in causal inference about the effect of marital status on the transition to home ownership, the findings would be biased.3

The sample sizes after optimal-pair and optimal full matching are also presented in table 1. After optimal-pair matching, the matched sample contains all 179 married participants and only 179 paired or matched unmarried participants. Note that the optimal-pair matching is estimated to retain all married participants and does not lose any married subject. Optimal full matching is estimated to retain all 179 married and 374 unmarried participants. It groups these participants in matched strata. The ratio of treated (married) participants to control (unmarried) participants varies by stratum, but the married participants within each stratum are estimated to share a propensity score similar to that of the unmarried participants within the stratum. Note that the optimal full matching employs all subjects from the original sample.

As table 1 indicates, both optimal-pair matching and optimal full matching are estimated to improve sample balances. The absolute standardized difference in covariate means before optimal matching ($d_X$) generally has a higher value than the index after optimal matching ($d_{Xm}$). For example, the $d_X$ of gender before optimal matching is .304. This means that the treatment and control groups are 30.4 percent of a standard deviation apart on gender. After optimal-pair matching, the $d_{Xu}$ of gender for the same covariate is .054. In other words, the two
## Table 1

**Sample Description and Imbalance Check before and after Matching**

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Overall Sample before Matching (n = 553)</th>
<th>ASDCM after Optimal Matching (d_{std})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of Mean(^a)</td>
<td>Pair Matching (n = 358)</td>
</tr>
<tr>
<td>No. married</td>
<td>179</td>
<td>179</td>
</tr>
<tr>
<td>No. nonmarried</td>
<td>374</td>
<td>179</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>42.1(^**)</td>
<td>.304</td>
</tr>
<tr>
<td>Female</td>
<td>27.7(^**)</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>39.2(^***)</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>20.4(^***)</td>
<td>.468(^b)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>42.7(^***)</td>
<td>.241(^b)</td>
</tr>
<tr>
<td>Other</td>
<td>40.0(^***)</td>
<td>.012</td>
</tr>
<tr>
<td>Age at baseline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>35.9 (11.3)</td>
<td>.114</td>
</tr>
<tr>
<td>Nonmarried</td>
<td>36.0 (11.3)</td>
<td></td>
</tr>
<tr>
<td>Education at baseline(^c)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>3.41 (2.08)</td>
<td>.132</td>
</tr>
<tr>
<td>Nonmarried</td>
<td>3.18 (1.96)</td>
<td></td>
</tr>
<tr>
<td>No. of children at baseline(^c)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>.81 (1.15)</td>
<td>.130</td>
</tr>
<tr>
<td>Nonmarried</td>
<td>.67 (1.03)</td>
<td></td>
</tr>
<tr>
<td>Employment status at baseline(^c)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working</td>
<td>34.3</td>
<td>.480</td>
</tr>
<tr>
<td>Not working</td>
<td>28.3</td>
<td></td>
</tr>
<tr>
<td>Income at baseline (in $1,000s)(^c)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>25.24(^***) (13.05)</td>
<td>.088</td>
</tr>
<tr>
<td>Nonmarried</td>
<td>19.24(^***) (11.92)</td>
<td></td>
</tr>
<tr>
<td>Census tract’s median house value</td>
<td></td>
<td>.110</td>
</tr>
<tr>
<td>Married</td>
<td>90,392.7 (36,052.5)</td>
<td></td>
</tr>
<tr>
<td>Nonmarried</td>
<td>94,590.1 (39,867.2)</td>
<td></td>
</tr>
<tr>
<td>Census tract’s median rent value</td>
<td></td>
<td>.088</td>
</tr>
<tr>
<td>Married</td>
<td>479.80 (125.60)</td>
<td></td>
</tr>
<tr>
<td>Nonmarried</td>
<td>468.49 (131.62)</td>
<td></td>
</tr>
<tr>
<td>Census tract’s disadvantage score</td>
<td></td>
<td>.180</td>
</tr>
<tr>
<td>Married</td>
<td>.13 (.57)</td>
<td></td>
</tr>
<tr>
<td>Nonmarried</td>
<td>.24 (.67)</td>
<td></td>
</tr>
</tbody>
</table>

Note.—ASDCM = absolute standardized difference in covariate means.

\(^a\) Each entry is the percentage of home owners in the categorical covariate or the mean of the continuous covariate by group. Standard deviations are presented in parentheses.

\(^b\) Race is recorded as two dummy variables: African American; Hispanic and others. White serves as a reference category.

\(^c\) Measured as an ordinal variable.

\(^**\) p < .01.
\(^***\) p < .001.
groups differ on gender by 5.4 percent of a standard deviation. After optimal full matching, the \( d_{Xm} \) of gender is \(.084\); the two groups differ on gender by \(8.4\) percent of a standard deviation. The value of most covariates decreases from the \( d_X \) to the \( d_{Xm} \), suggesting that optimal matching indeed improves balances. There is only one covariate (age at baseline) in which the \( d_{Xm} \) for pair matching and that for full matching are greater than their corresponding \( d_X \); since the two groups differ after matching on age at baseline by about \(11\) percent of a standard deviation, the sample can be considered balanced.

When conducting propensity-score matching, it is important to take into account whether the treatment (married) and control (unmarried) groups have a similar distribution on propensity scores. Figures 1 and 2 show the histograms and box plots, respectively, of the estimated propensity scores derived using GBM. As figure 1 indicates, the two groups differ on the distribution of estimated propensity scores, sharing a very narrow common-support region. This narrow region of common support is especially problematic because, if one applies nearest-neighbor matching within caliper or other types of greedy matching, the narrow common-support region results in a great loss of matched participants. However, this is not a problem if one uses optimal matching;
both pair matching and full matching create matches for all 179 married participants. Further, the reduction of the sample size occurs only among unmarried participants in the pair matching. By design, each participant matches only one control.

Table 2 presents results of the survival analysis. For this study, the findings are consistent across all three models that estimate differences between married and unmarried participants in the timing of a house purchase. For the original unmatched sample, the estimated odds ratio of purchasing a home is 3.728 ($p < .001$). This indicates that the odds of home ownership are 272.8 percent higher for married participants than for unmarried ones (i.e., \([3.728 - 1] \times 100 = 272.8\) percent). Similarly, the odds ratio of a house purchase for the matched sample created by the optimal-pair matching is 3.445 (244.5 percent higher for married participants; $p < .001$). Both findings suggest that married participants purchased houses during the study period at a faster rate than the unmarried participants did.

In the optimal, full, matched sample, the length of time that married participants take to purchase a house is estimated to be .308 of a year shorter (approximately 3.70 months shorter) than the time taken by
unmarried participants. The Hodges-Lehmann aligned rank test shows that this is a statistically significant difference ($p < .05$).

Other statistically significant predictors of the timing of a house purchase include the participant’s age at baseline ($p < .01$ for the overall sample), educational attainment ($p < .01$ for the overall and matched samples), and income ($p < .001$ for the overall and matched samples). In general, participants who are younger, attained higher levels of education, and have higher income are also likely to purchase houses at a faster time-to-event rate.

The first sensitivity analysis indicates the robustness of the current findings. Using a sample that broadly defines nonmarried people to include those who cohabitated, divorced, and are widowed ($N = 923$, of whom 224 were married at some point and 699 were never married),

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Estimated Odds Ratio from Discrete-Time Model</th>
<th>Mean Difference after Matching ($n = 553$)*a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Before Matching ($n = 553$)</td>
<td>After Optimal-Pair Matching ($n = 358$)</td>
</tr>
<tr>
<td>Marital status (nonmarried):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>3.728***</td>
<td>3.445***</td>
</tr>
<tr>
<td>Gender (female):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>.999</td>
<td>.980</td>
</tr>
<tr>
<td>Race (white):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>1.119</td>
<td>1.050</td>
</tr>
<tr>
<td>Hispanic or other</td>
<td>1.201</td>
<td>1.243</td>
</tr>
<tr>
<td>Age at baseline</td>
<td>.974**</td>
<td>.980</td>
</tr>
<tr>
<td>Education at baseline</td>
<td>1.149**</td>
<td>1.178**</td>
</tr>
<tr>
<td>Year indicator variable (year 4):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 1</td>
<td>1.509</td>
<td>2.001</td>
</tr>
<tr>
<td>Year 2</td>
<td>2.321**</td>
<td>3.058**</td>
</tr>
<tr>
<td>Year 3</td>
<td>1.265</td>
<td>1.818</td>
</tr>
</tbody>
</table>

Note.—Reference group is shown in parentheses for the categorical covariate.

$a$ Mean difference of time to event with the Hodges-Lehmann aligned rank test for the sample after optimal full matching.

$b$ One-tailed test based on hypothesized negative direction.

c Measured as an ordinal variable.

d Time-varying covariate.

* $p < .05$.

** $p < .01$.

*** $p < .001$. 

**Table 2**

**Results of the Discrete-Time Models and the Hodges-Lehmann Aligned Rank Test**
the discrete-time model shows that the odds ratio of purchasing a house within 4 years is 194.8 percent higher for the married group than for the unmarried group ($p < .001$). The discrete-time model based on the optimal-pair matching shows that odds for the married group are 157.9 percent higher than for the unmarried group ($p < .001$). Estimates in the optimal full matching model show that the length of time married people take to purchase a house is .373 of a year shorter (approximately 4.48 months shorter) than the time unmarried people take ($p < .05$). Note that defining the unmarried category more broadly is estimated to reduce the magnitude of the effect of marriage on the timing of purchasing a house (or to bias the findings downward), but the result remains statistically significant. This is consistent with the authors’ speculation about those other groups, the cohabiting, divorced, and widowed participants, who are somewhat less likely to be at a stage in life when they can purchase a home.

The second sensitivity analysis, which employs greedy matching, reveals the same findings as those obtained through optimal matching; although greedy matching excludes cases that fall apart from the common-support region, the results still suggest that the rate at which married people purchase houses is faster than that for purchases by non-married people ($p < .001$). As the authors expected, excluding cases that fall outside the common-support region strengthens the relationship between marriage and the timing of a house purchase. For example, in the results based on the broader definition of unmarried people ($N = 923$), the odds ratio estimated through optimal-pair matching is 2.579 ($p < .001$), but that estimated through greedy matching is 2.711 ($p < .001$).

In summary, this study generally confirms the research hypothesis and finds that marriage has a positive effect on home ownership. This conclusion gives some validity to models that include important covariates affecting the selection between getting married and not getting married. These covariates are controlled through a series of propensity-score matching and analytic approaches.

Discussion

Using 5 years of data from CAPS, this study addresses the question of whether marital status has a net and strong correlation with the frequency and speed of transitions into home ownership. In results from models that use propensity-score matching to control for selection bias between married and unmarried groups, low- and moderate-income married couples are found to have higher odds of buying a home and to purchase homes at faster rates than their nonmarried counterparts.

These findings support the theoretical framework of this research. Economic and sociological perspectives suggest that married individuals
are more apt to purchase a home because they face fewer borrowing constraints (Linneman and Wachter 1989). They also may have attributes that make them more likely to own a home than are single householders (Lupton and Smith 2003). Further, unmarried households may be less likely to buy a home than are married counterparts because of the transitory nature of unmarried household members’ lives and because unmarried households may lack their counterparts’ need for large living quarters (Hendershott et al. 2009).

One reason that married couples are more likely than unmarried householders to buy homes is that married couples have more potential wage earners within the home. However, there are other factors at work as well. First, individuals in American society face strong normative pressures to organize life-course events in specific patterns. There is an expectation that marriage precedes home purchase in this sequence (Townsend 2002). Individuals respond to this social pressure by shaping their expectations, aspirations, and behavior to adhere to these societal norms.

Institutional and environmental factors also may condition the ability of unmarried people to purchase homes relative to that of married couples. For example, cohabiting couples may have more difficulty obtaining mortgage loans or may receive less attractive loan terms than married couples. Because the risk of relationship dissolution is higher among unmarried couples than among their married counterparts, mainline banking institutions may perceive unmarried couples as less creditworthy and, fearing mortgage default, may be hesitant to offer them financing for home ownership (Leppel 2007). The move to automated underwriting decisions reduces the role of social norms in lending decisions. Beyond the measured individual characteristics, however, these forces may still affect the likelihood of home ownership for unmarried persons relative to that for married couples.

It is equally important to consider what aspects of the marital relationship drive this study’s findings. Legal proceedings define marriage, but few would suggest that those proceedings make a couple more likely to buy a home. The authors instead suggest that the changing social status and life-course position of the relationship alter the social norms faced by the couple and their expectations for life in the future. In addition, common attributes of marriage, such as stability and commitment, may be crucial determinants that explain these results. The authors note that the results are consistent with an existing economic literature, which emphasizes the risk sharing associated with marriage (Mazzocco 2004).

A limitation of this study is worth acknowledging. Propensity-score matching fails to correct for selection bias due to unmeasured variables. Unlike randomized clinical trials that balance data on both measured and unmeasured variables, propensity-score matching cannot correct
for hidden selection bias (Guo and Fraser 2010). Thus, if the matching process omits important variables that predict marriage, the study findings will be prone to error. The consistency of these findings across all the samples, matched and unmatched, is not uncommon in results from observational studies in which the cause has a strong impact on effect. Although the matching may miss important covariates that affect sample selection, the results of the analysis are consistent and revealing.

The ideal study would include rich, detailed, individual-level predictors of marriage, such as those used by Ronald Mincy, Jennifer Hill, and Marilyn Sinkewicz (2009), but such a wide array of characteristics is not available in this study’s data set. Nonetheless, the propensity-scoring model in this study includes all available variables theorized to predict the marital status of respondents. The model also includes variables that may confound the estimated relationship between marital status and the home ownership choice. The authors considered brute force approaches to generating covariates in the propensity-score model (such as stepwise regression and data mining) but ultimately rejected these approaches as atheoretical, instead choosing a model grounded in the literature on the predictors of marital choice. In developing the model to estimate the propensity score, the authors tested other specifications and the inclusion of additional covariates. Those other iterations are found to be inferior to the model presented in this article.

Because propensity-score analysis fails to balance study conditions that are due to unmeasured variables, this study cannot claim that the relationship between marriage and the transition into home ownership is causal. However, the analytic methods in this study are carefully chosen to overcome limitations of the conventional covariance control approach and therefore permit the authors to conclude that there is a likely net association between marriage and home ownership.

This study makes several contributions to the literature. First, it speaks directly to ongoing policy debates about the costs and benefits of promoting home ownership for low-income households. It is important that policies aimed at promoting home ownership recognize the close coupling of marriage and home ownership. Policies and programs that work for low-income married couples might not meet the needs of single people looking to become home owners. Alternatively, home ownership may not be the optimal asset-building strategy for some single people, and policy makers should also emphasize efforts to improve the rental experience for people not ready to shift to home ownership. As both policy makers and practitioners consider the role that home ownership should play in antipoverty efforts, it is important to understand who benefits from home ownership and who may not.

Second, the findings have important implications for asset-building policy and marriage promotion initiatives. The past decade has witnessed an increasing number of calls to promote marriage as a poverty
Marriage and Home Ownership

reduction strategy. However, critics counter that the benefits of marriage are oversold and that marriage promotion efforts do not translate to reductions in poverty (Alternatives to Marriage Project 2007). Further, women’s increasing economic independence and access to employment opportunities are cited as evidence that promoting marriage is an outdated approach to asset building (Seefeldt and Smock 2004). However, this study’s findings suggest that policy makers may want to reexamine the possibility that marriage and home ownership have a combined effect on the economic well-being of low-income households.

Finally, this study offers one of the first attempts to use propensity-score analysis to address self-selection into marriage. Most studies on the effect of marriage fall short because they fail to model the choice individuals make to be married and how that choice affects the outcome of interest. Instead of the conventional covariance control approach, which cannot adjust for selection bias, this study uses an innovative and rigorous method that allows the authors to draw causal inference between marital status and the transition to home ownership.

Although this study focuses specifically on home ownership rather than on asset building in general, the authors nonetheless conclude that one benefit of marriage is its function as a catalyst to home ownership for low-income families. Previous studies provide strong evidence that, if a home is purchased as a long-term investment, responsible home ownership can yield meaningful wealth returns, even for low-income families (Riley and Quercia 2011). Given the social and economic gains that may be gleaned from marriage, and in turn from home ownership, policy makers may want to formulate well-coordinated policies aimed at concurrently increasing marriage and home ownership opportunities. Such policies could substantially benefit disadvantaged families that struggle to achieve economic and familial stability.

Appendix

Propensity-Score Optimal Matching

As discussed by Shenyang Guo and Mark Fraser (2010), a greedy matching method has several limitations. In dividing matching into a series of discreet decisions, it fails to account for the effect of a given match on the overall efficiency of matching. It can sometimes produce too many unmatched cases or too many inexactly matched cases. Finally, it requires a sizable region of common support to work efficiently. To overcome these limitations, this study applies the optimal-matching method (Rosenbaum 2002).

The optimal-matching method uses the network flow theory to optimize the creation of a matched sample. A primary feature of network flow is that it concerns the cost of using \( b \) for \( a \) as a match, where a
cost is defined as the effect of having the pair \((a, b)\) on the total distance of propensity scores.

Initially, there are two sets of participants: the treated participants are in a set \(A\), and the controls are in a set \(B\), with \(A \cap B = \emptyset\). The initial number of treated participants is \(|A|\), and the number of controls is \(|B|\), where \(|\cdot|\) denotes the number of elements of a set.

For each \(a \in A\) and each \(b \in B\), there is a distance, \(\delta_{ab}\), with \(0 \leq \delta_{ab} \leq \infty\). The distance measures the difference between \(a\) and \(b\) in terms of their observed covariates, such as their difference on propensity scores. Matching is a process to develop \(S\) strata \((A_1, \ldots, A_S; B_1, \ldots, B_S)\) consisting of \(S\) nonempty, disjoint participants of \(A\) and \(S\) nonempty, disjoint subsets of \(B\), so that \(|A_1| \geq 1\), \(|B_1| \geq 1\), \(A_i \cap A_j = \emptyset\) for \(s \neq s'\), \(B_i \cap B_j = \emptyset\) for \(s \neq s'\), \(A_1 \cup \ldots \cup A_S \subseteq A\), and \(B_1 \cup \ldots \cup B_S \subseteq B\).

By this definition, a matching process produces \(S\) matched sets, each of which contains \(|A_1|\) and \(|B_1|\), \(|A_2|\) and \(|B_2|\), \ldots, \(|A_S|\) and \(|B_S|\). Notice that, by definition, within a stratum or matched set, treated participants are similar to controls in terms of propensity scores. Depending on the structure (i.e., the ratio of the number of treated participants to the number of control participants within each stratum) the analyst imposes on matching, matching may be classified into the following three types: pair matching, matching using a variable ratio, and full matching. Optimal matching is the process of developing matched sets \((A_1, \ldots, A_S; B_1, \ldots, B_S)\) with size \((\alpha, \beta)\) in such a way that the total sample distance of propensity scores is minimized. Formally, optimal matching minimizes the total distance, \(\Delta\), defined as:

\[
\Delta = \sum_{s=1}^{S} \omega(|A_i|, |B_i|)\delta(A_i, B_i),
\]

where \(\omega(|A_i|, |B_i|)\) is a weight function.

Using the R program optmatch (Hansen 2007) and following the guidelines suggested by Paul Rosenbaum (2002), this study employs optimal-pair matching and optimal full matching. Using the matched sample created by optimal-pair matching, the authors conducted the discrete-time survival analysis as previously described. Next, the matched sample created by optimal full matching was used to conduct the Hodges-Lehmann aligned rank test, which is described below.

**Generalized Boosted Modeling**

This study uses GBM regression to estimate the propensity scores (McCaffrey, Ridgeway, and Morral 2004). A program developed by Matthias Schonlau (2005) is used to estimate the GBM. One of the problems with the binary logistic regression lies in specifying an unknown functional form for each predictor. If specifying functional forms can be avoided, the search for the best model may lead to a more accurate
prediction of treatment probability. Regression trees are employed in GBM to estimate regression coefficients. A useful property of trees is that they are invariant to one-to-one transformations of the independent variables. Thus, “whether we use age, log(age), or age^2 as a participant’s attribute, we get exactly the same propensity score adjustments” (McCaffrey et al. 2004, 408).

The GBM regression employs an automated, data-adaptive algorithm that fits several models by way of a regression tree and then merges the predictions produced by each model. As such, GBM can be used with a large number of pretreatment covariates to fit a nonlinear surface and predict treatment assignment. It is one of the latest prediction methods to be adopted by the machine-learning community as well as by mainstream statistics research (Guo and Fraser 2010). From a statistical perspective, the breakthrough in applying boosting to logistic regression and exponential family models was made by Jerome Friedman, Trevor Hastie, and Robert Tibshirani (2000).

Checking Covariate Imbalance

An important task for propensity-score matching is to check covariate imbalance before and after matching. For most study covariates, the analyst hopes to achieve balance between treatment and control groups through matching. This study employs chi-square tests and independent sample t-tests to check covariate imbalance before and after greedy matching. It also uses the imbalance indexes developed by Amelia Haviland, Daniel Nagin, and Paul Rosenbaum (2007) to check covariate imbalance before and after optimal matching. The current study conducts this analysis with IMBALANCE, a Stata program developed by Guo (2008b). Two statistics, d_X and d_Xm, are generated by the imbalance check. They can be interpreted as the difference between treatment and control groups on X in terms of a standard deviation unit of X. Imbalance on X before matching is indicated by d_X, and d_Xm indicates imbalance on X after matching. Typically, the analyst expects to have d_X > d_Xm because she or he expects the need to correct for imbalance before matching, and the sample balance improves after matching.

The Hodges-Lehmann Aligned Rank Test

The outcome analysis following optimal full matching is complicated because the survival model analyzing event times for a matched sample needs to consider correlated survival times within matched stratum. Ideally, the analyst would use a frailty model that includes random effects to represent extra heterogeneity of the unit, heterogeneity that gives rise to the dependence of event times (Hougaard 2000; Kalbfleisch and
Although a frailty model is fruitful in matched studies with a randomized experiment, that approach has not yet been developed for such observational research as the current study. Thus, the Hodges-Lehmann aligned rank test (Hodges and Lehmann 1962) is used to evaluate outcome differences between treatment and control participants for the sample generated by the optimal full match.

The aim of the Hodges-Lehmann aligned rank test is to produce a crude estimate of the difference on the time-to-event data between study groups and to gauge whether that difference is statistically significant. However, the Hodges-Lehmann approach has three inherent limitations: (a) it treats the length of time to the event as uncensored, (b) it analyzes mean difference rather than differences of other statistics that better account for skewed distribution of survival times, and (c) it is bivariate and fails to control for covariates of the survival times.

This study employs the Hodges-Lehmann method to evaluate the sample-average treatment effect by assessing a weighted average of the mean differences between treatment and control participants of all matched sets. This is represented as

$$\hat{\delta} = \sum_{i=1}^{b} \frac{n_i + m_i}{N} (\bar{T}_{0i} - \bar{T}_{1i}),$$

where $i$ indexes the $b$ matched strata, $N$ is the total number of sample participants, $n_i$ is the number of treated participants in the $i$th stratum, $m_i$ is the number of controls in the $i$th stratum, and $\bar{T}_{0i}$ and $\bar{T}_{1i}$ are the mean survival times corresponding to the control and treatment groups in the $i$th stratum.

The Hodges-Lehmann aligned rank test (Hodges and Lehmann 1962; Lehmann 2006) further assesses whether this average treatment effect differs to a statistically significant degree. To conduct this analysis, the study employs hodgesl, a Stata program developed by Guo (2008a).

**Propensity-Score Greedy Matching**

This study employs the propensity-score greedy-matching technique (Rosenbaum and Rubin 1983, 1985) as a sensitivity analysis. It involves the following steps.

First, it uses the binary logistic regression to estimate a propensity score of receiving treatment (i.e., being married). By definition, a propensity score is a conditional probability of a participant receiving treatment given observed covariates. The advantage of using propensity-score matching is that it reduces dimensions; the conditioning variables the study aims to match may include many covariates. The propensity-score approach reduces all this dimensionality to a one-dimensional score. In doing so, the approach eases the burden of finding matches within the study sample. Following Rosenbaum and Rubin (1985), the study em-
ploys the logit of the predicted probability from the logistic regression as a propensity score. That is,

\[
\hat{q}(x) = \log \left( \frac{1 - \hat{e}(x)}{\hat{e}(x)} \right),
\]

where \( \hat{e}(x) \) is the predicted probability from the logistic regression because the distribution of \( \hat{q}(x) \) approaches a normal distribution. The logistic regression employs the same set of independent variables as those used in the discrete-time model, except that the three time-varying variables (i.e., participant’s employment status, number of children, and income) are specified as time-fixed variables and measured at the baseline interview.

Second, the propensity-score approach matches treatment group participants with control group participants on estimated propensity scores to strengthen the validity of the estimate of counterfactuals (i.e., outcome values of the comparison group). This study employs the nearest neighbor matching within a caliper (Rosenbaum and Rubin 1985). The method selects a control participant, \( j \), as a match for treatment participant, \( i \), if and only if the absolute distance between the respective propensity scores for the two participants (i.e., the difference between propensity scores \( P_i \) and \( P_j \)) meets the following condition:

\[
P_i - P_j < \varepsilon,
\]

where \( \varepsilon \) is a prespecified tolerance for matching, or a caliper. Rosenbaum and Rubin (1985) suggest using a caliper size of one-quarter of a standard deviation of the sample-estimated propensity scores (i.e., \( \varepsilon \leq .25\sigma_P \), where \( \sigma_P \) denotes the standard deviation of the estimated propensity scores of the sample).

Finally, the study conducts the discrete-time model with the matched sample to study outcome differences between treatment and control participants. As mentioned earlier, this analysis is expected to provide a more valid estimate of causal effect than that obtained from other types of analysis because it uses a sophisticated control of selection bias.

Notes

The authors thank the National Center for Family and Marriage Research and the Ford Foundation for their support of this research. They also thank Janneke Ratcliffe, with the Center for Community Capital, and Laurie Graham and Krista Holub, from the School of Social Work, for their valuable comments and suggestions.

1. When eligible renters could not be found within the census block, researchers expanded the radius of the search area up to 4 miles.

2. Other work published from these data uses variables that appear to be related to this investigation. Because of the modular nature of the CAPS data, these variables are unavailable in all of the years for the population under study here. Thus, they could not be included in the current study’s models.

3. Results of the propensity-score model using GBM are not presented due to space limitation but are available on request.
References


Grinstein-Weiss, Michal, Min Zhan, and Michael Sherraden. 2006. “Saving Performance...


