Title
Hazard Models of Changing Household Demographics

Author
Kazimi, Camilla

Publication Date
1994-12-01

Hazard Models of Changing Household Demographics

UCI-ITS-WP-94-11

Camilla Kazimi

Department of Economics and
Institute of Transportation Studies
University of California, Irvine

December 1994

Institute of Transportation Studies
University of California, Irvine
Irvine, CA 92697-3600, U.S.A.
http://www.its.uci.edu

Prepared for the 74th Annual Meeting of the Transportation Research Board, Washington, D.C., January 1995.

Hazard Models of Changing Household Demographics

Camilla Kazimi

Abstract

In this paper, I develop demographic models which can be used to simulation household changes resulting from marriage, divorce or separation, childbirth, children leaving home, cohabitation, extended families living together, death, and so forth. They are dynamic in nature, and are meant to be used within a larger microsimulation system. In fact, they can be used by any microsimulation system that models decision-making at the household level. They extend previous work in three ways: 1) by using continuos time hazard models, 2) by allowing for inter-dependencies across the various type of changes that a household may undergo, and 3) by including several important covariates. These covariates include age, gender, race, education, income, employment status, and indicators for previous demographic events (e.g. birth of a child out-of-wedlock and previous marriages). They provide insight into the demographic patterns across different socioeconomic groups.

I am very grateful to David Brownstone for his guidance and motivation on this project, and to Kenneth Small for his detailed comments. Portions of this project were funded by Southern California Edison Co. and Pacific Gas and Electric Co. through the Institute of Transportation Studies. Additional funding is being provided by the U. S. Department of Transportation and the California Department of Transportation (UCTC graduate dissertation fellowship). I am solely responsible for the paper's content.

1. Introduction

Dynamic microsimulation models are used to determine the impact of economic policy and technological change. They answer "what if" questions such as what will be the effect of a new income tax policy, or what will happen if the current welfare program is cut back (Orcutt et al., 1986). In the transportation field, they have been used to predict future vehicle demand and the resulting infrastructure requirements (Hensher et al, 1992). When stated-preference information is available, they can be used to predict consumers' adoption of new technology.

Currently the Institute of Transportation Studies (ITS) at the University of California is developing a modeling system which simulates vehicle demand and usage as alternative-fuel vehicles are introduced into the market. Mandates in California require the sale of zero-emission vehicles (effectively electric vehicles given current technology): 2% of automobile sales by major manufacturers must be zero-emission vehicles in 1998, 5% in 2001, and 10% in 2003. In addition, other alternative-fuel vehicles such as compressed natural gas, methanol, and various hybrids may be required to meet stricter emission standards. These mandates and emission standards are quite controversial: auto manufacturers maintain that consumers will not purchase electric vehicles because conventional gasoline automobiles will dominate in terms of price and performance. The bottom line is that simply mandating sales does not guarantee purchase and usage of lower-emitting vehicles. Questions still remain: Under reasonable technological assumption, what will the demand for alternative-fuel vehicles will be? Will they replace older polluting cars or newer cleaner ones? If purchased, how will they be used? The microsimulation models will address these questions.

In general, microsimulation models begin with some sample of individuals or households from the population. Each period the sample members are faced with changing circumstances (such as the introduction of a new vehicle type), and their responses are forecast based on models of their decision-making process. The models are dynamic when current decisions affect the options available next year, and so forth throughout the time frame of the study.

For vehicle choice and usage models, changes in household structure will almost certainly affect the outcome of decision-making process. The automobile is a large consumer durable that must meet the needs of the entire family. A thirty-year-old single man has different transportation needs than a thirty-year-old married man with two young children, and his automotive choice is likely to differ as well. Following common stereotypes, the single man drives a sports car while the married man drives a station wagon. The original sample of households will certainly undergo a series of demographic changes during the period of the microsimulation. People will marry, divorce and separate, have children, and so forth.

The purpose of this paper is to develop demographic models that will be able to simulate these changes. They are dynamic in nature, and can be used by other microsimulation systems that model decision-making at the household level. They extend previous work in three ways: 1) by using continuous time hazard models, 2) by allowing for inter-dependencies across the various types of changes that a household may undergo, and 3) by including several explanatory variables.

These demographic models are interesting in their own right. They provide insight into different demographic patterns across socioeconomic groups. For example, I find that single black women are more likely to have a child out-of-wedlock than their white counterparts, all else being equal. And the differences are quite dramatic. They are also less likely to marry. White women with a higher education and income levels are less likely to have a child out-of wedlock and less likely to marry early than their white counterparts with less education and income. Differences also exist between first and second (or higher) marriages. Holding all other factors constants (such as age, race, income, education, and gender), individuals that have previously been married are likely to remarry sooner than individuals who have not married for the first time.

These are just a few of the important differences. Several more are given in the estimation results in section 6.1, and illustrated by the survivor curves in section 6.1.1.

2 Changing Demographics in the United States

Households and families in the United States have undergone radical demographic changes over the past 30 to 40 years. In the 1950's, most people lived either with their parents or in college housing (supported by their parents) until they married. After marriage, women typically stayed home and raised children while their husbands went to work. Since then women have entered the labor force in large numbers. Women often experience a period of living on their own, working and independently supporting themselves. In 1965, 38.1% of all white women (regardless of their marital status) and 48.6% of nonwhite women were employed. By 1984 these numbers had grown to 53.3% and 55.% respectively (Blau and Ferber, 1986). The increases are even more dramatic when we break down the figures by marital status.

Table 1
Labor Force Participation Rates for Women

	1966	1984
Never married	40.8	63.3
Married, husband present	35.4	52.8
Other (married at one time)	39.5	44.9
Married, husband absent	n.a.	61.1
Widowed	n.a.	20.4
Divorced	n.a.	74.3

Source: US. Department of Labor, Bureau of Labor Statistics, Special Labor Force Report, no. 2163, Table B-5, p. 16 and Bureau of Labor Statistics data reported in Bureau of National Affairs, Daily Labor Report, no. 145 (July 27, 1984), p. B-3.

At the same time that women were entering the labor force, families was undergoing tremendous change. Divorce was becoming more prevalent, increasing numbers of children were born to unwed mothers, many were postponing childbearing or choosing to forego the experience altogether, and cohabitation became more common. The following subsections present and discuss data on each of these trends.

2.1 Divorce and Separation

Compared to other countries, the United States has a high divorce rate. In 1976, 5 out of every 1000 people divorced. Other countries with relatively high divorce rates in 1976 included Australia at 4.3 (per 1000), USSR at 3.4, Sweden at 2.7, Denmark at 2.5, Canada at 2.2, Egypt at 2.2, and Finland at 2.1. The median time between first marriage and divorce was approximately 7 years in the United States, between divorce and remarriage was 3 years, and between remarriage and second divorce was 5 years for those that passed through each phase. (Glick and Norton, 1977). Since most data sources do not distinguish between divorce and separation, marital disruption will refer to an occurrence of either event.

Using life table estimates based on the 1985 Current Population Survey, Martin and Bumpass (1989) have projected that nearly two-thirds of all marriages in the late 1980's will end in divorce or separation. Others argue that the rate is lower. While the marital disruption rate rose dramatically from the late 1960's to 1970's, it has declined slightly during the 1980's so that roughly half of all current marriages with partners in their thirties are likely to dissolve (Glick, 1990).

The disruption rate for first marriages varies by several factors. As shown in Table 2, disruption rates within the first 5 years of marriage have increased across all categories from the period 1970-1974 to the period 1980-1985. Those who marry at earlier ages, have less education, have children before marriage, or are of African American ethnicity are more likely to experience marital disruption within the first 5 years. Even after accounting for differences in education, employment status, and premarital births, African Americans are still at higher risk.

Table 2: Proportion of First Marriages Disrupted Within 5 Years

	1970-74	1980-85
Age at Marriage		
14-19	0.23	0.31
20-22	0.14	0.26
23-29	0.11	0.15
30+	0.14	0.14
Education		
0-11 years	0.21	0.33
High school graduate	0.18	0.26
Some college (or beyond)	0.16	0.16
Kids before marriage		
0	0.17	0.21
1+	0.22	0.36
Race		
White	0.17	0.22
African American	0.24	0.36
Hispanic	0.15	0.24

Source: Martin. T. C. and L. L. Bumpass (1989), Table 1, p. 41.

Others find similar results (Glick and Norton, 1977; Spanier and Glick, 1981; Bennet, et al., 1989; Heaton and Jacobson, 1994). While the previous table did not separately identify education beyond college, Glick and Norton (1977) find that women with advanced education (17 or more years of schooling) are more likely than women with high school or college degrees to end their marriage.

Using two-state hazard models and data from the 1982 and 1988 National Surveys of Family Growth , Heaton and Jacobson (1994) find that about half of black marriages and a third of white marriages will have ended within 15 years of marriage. Age at marriage has a large negative effect on divorce for whites (e.g. marrying at a young age increases the chance of marital disruption), but virtually no effect for blacks. Racial differences persist even after accounting for differences in mother's education, religion, region, and age at marriage.

Along another vein, families with sons are more likely to stay together than families with daughters. "Sons reduce the risk of marital disruption by 9% more than do daughters. The differences hold across marriage cohorts, racial groups, and categories of mother's education."[1] Fathers may be more involved in the upbringing of their sons and therefore more committed to the marriage. In addition, families with only 1 child are more likely break apart as compared to families with 2 or more children (Morgan, et al., Figure 2, p. 118).

Marital disruption rates for second marriages follow similar patterns as first marriages. Table 3 shows that rates in general are higher for second marriages as compared to first marriages (see previous table). Nonetheless, similar factors are associated with higher rates of separation and divorce. Those who marry at earlier ages, are less educated, and are African American are more likely to divorce or separate within the first 5 years. Children before marriage (presumably from the previous marriage) are not as detrimental as children born out-of-wedlock before the first marriage.

Table 3: Proportion of Second Marriages Disrupted Within 5 Years

	1970-74	1980-85
Age at Second Marriage		
14-19	0.26	0.40
20-22	0.15	0.26
23-29	0.17	0.27
30+	0.13	0.14
Education		
0-11 years	0.17	0.36
High school graduate	0.17	0.26
Some college (or beyond)	0.20	0.22
Kids before marriage		
0	0.16	0.24
1+	0.18	0.28
Race		
White	0.18	0.26
African American	0.21	0.43
Hispanic	0.10	0.28

Source: Martin. T. C. and L. L. Bumpass (1989). Table 3, p. 45.

[1] Morgan. S. P.. Lye. D. N.. and G. A. Condran (1988). p. 110.

There are important areas of conflict within families which are not captured by simple measures such as race, educational background, age at marriage, and so forth. Some argue that marriages entail a balance of power between husband and wife. Events which tip the balance without the consent of both parties causes stress, potentially ending in divorce. One common event which may cause such stress occurs when the wife works outside the home. Husbands may view such an arrangement as either positive or negative. Pyke (1994) argues that it may be viewed negatively when the "husband suffers low occupational status or chronic unemployment" in which case the husband will "devalue their wife's market work and view it as a burden (p. 75)." In addition, husbands may feel inadequate as the breadwinner in the family when the wife earns more (Stanley, et al. 1986). Men in general may find it difficult to relate to women who earn more (Bane1976), and empathy is often an important component of marital stability.

2.2 Marriage

The United States appears to have similar marriage rates when compared to other industrialized nations. In 1976, the marriage rate in the United States (defined as the number of marriages divided by the population in thousands) was 9.9. Corresponding rates were 8.1 in Australia, 8.7 in Canada, 10.0 in Egypt, 8.5 in Israel, 7.8 in Japan, and 10.1 in the USSR (Glick, 1977).

But singles today are more likely to postpone marriage. For example, in 1960 approximately 11% of women in their early thirties were unmarried. By 1987 that figure had increased 2 1/2 times to 27% . It is estimated that around 10% of young adults in the 1980's were likely to remain unmarried throughout their entire life (Glick, 1990). An even more dramatic comparison can be made between the turn of the century and current time.

> "In 1890, the median age of the wife at marriage was 22 years and the median age when her husband died was only 53 years. (There was) a fifty-fifty chance that the

> marriage would actually end before the last child left home. Because of the much longer length of life today, the corresponding age of the wife at dissolution of the marriage is now 68 years." (Glick 1990)

Marriages used to last about 30 years, but today may last around 50 years assuming the couple does not divorce.

Within the United States, marriage rates vary by race. Heaton and Jacobson (1994) use two-state hazard models to examine first marriages. They find a dramatic difference between the marriage patterns of whites and African Americans. Fewer and fewer African Americans are choosing to marry. For example, 85% of white women will marry between the ages of 19 and 25, whereas less than 65% of African American women will be married by the age of 30. Bennett, et al. (1989) find similar differences, and argue that African American women are faced with a smaller pool of available mates. Several factors contribute to this fact including a smaller male to female birth ratio and higher death and incarceration rates for young black males.

The effects of education varies depending upon race. For whites, higher education (college and beyond) is negatively associated with the chance of ever marrying, but the association is weak. On the other hand, higher education is positive and significantly associated with the probability of ever marrying for blacks (Bennet, et al., 1989). In general, increased education is associated with delayed marriage (Heaton and Jacobson, 1994).

Remarriage rates follow slightly different patterns. First, young adults who have divorced are much more likely to remarry when compared to similar cohorts who have yet to marry (Glick, 1990). This may to due to sampling bias since those who have already married probably have a higher propensity to marry in general.

Remarriage rates vary by education and income as shown by the figures for 1980 in Table 4. Increased education is associated with lower remarriage rates. The differences are more pronounced for women, with rates dropping from 61% for high school dropouts to 42% for women with graduate work. Both men and women with college degrees are less likely to remarry because they face a small pool of eligible mates as a result of the fact that college-educated people are more likely to have stable first marriages. Men with higher income and women with lower income are most likely to remarry (not necessarily to each other).

Table 4: Remarriage Rates in 1980

	Men	Women
Education		
Some high school or less	0.67	0.61
High school graduate	0.67	0.56
Some college	0.65	0.50
College graduate	0.61	0.44
Graduate work or degree	0.59	0.42
Income		
< 8,000	0.54	0.58
8,000 - 15,000	0.65	0.45
15,000 - 25,000	0.69	0.38
> 25,000	0.70	0.39

Source: Glick, P. C. and S. Lin (1987), Tables 2 and 3, pp. 168-171.

Bumpass, Sweet, and Martin (1990) use proportional hazard models and data from the 1980 and 1985 Current Population Survey to identify differences in remarriage rates across different groups. They find that remarriage rates are 26% lower for women aged 30-39 at separation (63% lower for those over 40) as compared to those under 25. Women who were 22 year or older when they married for the first time have a 38% lower remarriage rate than those who married at a younger age for the first time. Women with children from their previous marriage have a 25% lower remarriage rate than those without children, supporting Becker's notion that children are capital specific to the marriage in which they were conceived (Becker, et al., 1977). Finally remarriage is much less common among blacks. Controlling for all other variables, they have a remarriage rate that is only 1/4 of their white counterparts.

2.3 Fertility and Child Rearing

The most dramatic changes with respect to fertility is the increasing numbers of children born to unwed mothers. In 1960 only 5% of births occurred to unwed mothers (2% of white births and 22% of black births). By 1987 25% of all births occurred to unmarried women (17% of white births and 62% of black births).[2] Coupled with high divorce and separation rates, this leads to the fact that 23% of families with children less than 18 years old were headed by a single parent in 1988.[3]

A second change over the past 40 years has been the reduction in family size. During the baby boom of the mid-1950's, the number of children born per woman was about 3.8. From 1970 on, the number has remained stable at 1.8 per woman.[4]

Race is strongly associated with the timing of childbearing. Black males and females are likely to be sexually active at earlier ages and more accepting of sexual activity at younger ages when compared to white cohorts (Cherlin, 1992; Moore and Steif, 1991). As a result, there is 25 point difference in the percentage of those who have given birth by age 21 between blacks and whites.

White women from 2-parent families with highly educated mothers are more likely to deter child birth. For black women, growing up in a 2-parent family has a much less significant impact while their mother's education is important in delaying child birth.[5]

2.4 Cohabitation

While more and more singles are delaying marriage, the incidence of cohabitation has been increasing. Nearly one third of all young unmarried adults (ages 18-35) will live together (Thornton, 1988).

---

[2] Glick, P. C. (1990), p. 139.
[3] Ibid, p. 141.
[4] Ibid, p. 140.
[5] Heaton and Jacobson (1994). The authors look only at the mother's educational influence on their daughter's fertility instead of the daughter's eduation.

Cohabitation is a substitute for marriage, and often actually precedes marriage. In the mid-1980's, nearly half of all first marriages were preceded by a cohabitation period (Bumpass, 1990). As the following table shows, the percent married has dropped dramatically from 1970 to 1985. At the same time, cohabitation rates have been increasing so that the percentage of adults in any form of union (marriage or cohabiting) has remained relatively stable

Table 5: Marriage and Cohabitation Proportions

	1970 % Ever Married	1985 % Ever Married	1985 % Ever in Union
Before Age 20			
Total	27	14	23
Males	18	5	13
Females	35	22	33
Whites	28	15	25
Blacks	24	6	15
< High school	43	30	49
High school graduate	42	17	28
College	14	5	10
Before Age 25			
Total	72	55	69
Males	65	43	59
Females	79	66	78
Whites	76	58	71
Blacks	61	37	61
< High school	72	58	76
High school graduate	80	62	74
College	66	48	62

Source: Bumpass, Sweet and Cherlin (1991), Table 1, p 916.

As might be expected, cohabitation is a less stable than marriage. The median length of cohabitation is 1.5 years (Bumpass, 1990). As far as differences among groups, there is a higher reported rate of cohabitation among men than women. Blacks are also more likely to cohabitate. In general, Blacks are more likely to live in all types of 'non-traditional' family arrangements (Heaton and Jacobson, 1994).

3 Demographic Modeling Strategies

Merz (1991) identifies two main types of dynamic microsimulation: longitudinal and cross-sectional. Both can be carried out at the household or individual level. Demographers typically focus on individual transitions while maintaining household records in accordance with individual-level changes. When actually implementing this strategy, the analyst must provide algorithms to maintain the household records.

Longitudinal microsimulation creates synthetic microunits (in this case, households) and forecasts their life cycle. Synthetic microunits are randomly assigned core characteristics based on the joint distribution of these features from a population sample. They are then assigned other non-core characteristics based on cross-tabulations of core and non-core features from the population sample. The synthetic households are weighted by the number of households that they represent in the population, and the weights are adjusted each period based on analyst-supplied forecasts of population growth and migration. The longitudinal technique has been used by Hensher et al. (1992) in their study of automobile demand and by Cowing and McFadden (1984) in their study of residential energy demand.

As an alternative method, cross-sectional microsimulation ages an actual population sample using empirically based transition probabilities. The number and composition of households changes from one period to the next, and the evolving demographics affect future behavior and transitions. Generally cross-sectional microsimulation is more difficult to apply than synthetic microunits, but the technique has several advantages. The analyst does not need to reweight the sample each period based on exogenous forecasts of population changes. In addition, a great deal of detail is maintained including the distributional impact of the policy under study. There are several ways to implement cross-sectional microsimulation including step-by-step submodules, multistate demography, and hazard models.

Step-by-step submodules have been used in several simulation programs including the Dynamic Simulation of Income Model (DYNASIM, 1976 or DYNASIM2, 1983 from the Urban Institute), MicroHaus (Gothenburg School of Economics, Sweden), and to a certain extent MIDAS (Goulis and Kitamura). Depending on the level of analysis, demographic changes occur as the individual or household sequentially passes through several submodules. For example, an individual becomes one year older, may marry, may divorce, may have a child, may become unemployed, and so on. Separate submodules are used for each demographic change, and the order the submodules is fixed. In other words, the individual first ages, then is subjected to the chance of marriage, then divorce and so forth through all the demographic processes. Often these systems do not account for interdependencies between transitions. As a result, changing the order of the submodules may change the outcome of the overall simulation.

Multistate demography (Land and Rogers, 1982; Rima and Van Wissen, 1987) determines the rate of movement between several analyst-defined states. A state describes the composition of and/or position within the household (e.g. head of a married couple, head of a family with 2 children, etc.). Movements from one state to another encompass several submodules from the previous step-by-step technique. The analyst must define several states which are typically at both the individual and household level. Assuming 'I' states at the individual level and 'H' states at the household level, the analyst determines movement rates across individual states (a matrix of I x I rates). These movement rates are then combined with fertility rates to determine movements across the household states (another H x H matrix). Some individual movements result in reclassification of the household, while others have no effect. The way in which individual movements affect households can become rather complex, and ground-rules must be determined in the design phase for handling actual household reclassifications (or movements across states). For example, a divorce may maintain the original household record after removing the husband and create a new household record containing only the husband.

The main disadvantage of multistate demography is the data requirements. If age is an important factor affecting individual movement rates, the analyst may define several age categories and the I x I matrix of movement rates must be determined for *each* age category. Extending the breakdown to include race and employment status would be unmanageable. It is unlikely that data would exist to fill in all the cells of the matrices of movement rates. If race and employment status are important explanatory variables of the those movement rates, simpler multistate demographic models which exclude those variables would produce suspect results. Nonetheless, multistate demography has theoretical grounding in the field of demographics and may be more consistent than tacking together several submodules.

An alternative to multistate demography uses hazard models. Hazard models measure the time until an individual or household undergoes some demographic change. In comparison to multistate demography, a hazard model may include several variables as determinants of then the change will occur instead of creating several transition matrices for each value of the variable. Therefore characteristics such as gender, ethnicity and income can easily be included in addition just age. For this reason I have chosen to use hazard models.

Previous work in this area has used discrete time hazard models (Davies, 1992). Davies also assumes that movement from the initial type of household to another is independent from the chance of movement from that initial type to any other household. This can be a very restrictive assumption, especially when important explanatory variables are excluded from the specification and estimation of the models.

My models will extend previous work in two ways: 1) by using continuous time hazard models and 2) by allowing for inter-dependencies across the movements from one household type to the various other possible types. The advantage of continuos time hazard models is that the hazard rate (defined as the instantaneous probability of movement from one household type to another given that the household has not made such a movement yet) can vary over any time interval. With discrete choice models, the hazard rate is assumed to be constant over the discrete time intervals. I will also allow for interdependencies between movements from the one type of household and other forms (referred to as unobserved heterogeneity below). In most cases, the movements rates are found to be independent.

3.1 Individual or Household Level Models?

What exactly is meant by the term 'household'? Households include individuals who live alone, families, cohabiting adults with or without children, and other extended families (for example, older parents who have moved in with their adult son or daughter's family). When defining a household, the litmus test is whether the group engages in shared consumption. Ermisch (1988, page 23) provides an economic definition of a household as "a unit which combines the time of its members and purchased goods and services in the production of outputs, at least some of which are shared among its members." The Census Bureau considers a group of individuals to be a household if they live in the same dwelling and share meals, and the alternative-fuel vehicle survey has adopted a similar definition (with the inclusion of sharing household expenses). Roommates would not typically be considered a household unit.

A longitudinal definition of a household can be rather ambiguous. What happens to a household when a couple divorces? Which person is a continuation of the original household? Duncan and Hill (1985) persuasively argue that simply restricting the analysis to households that remain intact will lead to biased results. For this reason, they argue that the analysis should be done at the individual level with the unit of measurement at the household level. In other words, individuals are described in terms of the type of household to which they belong. Individuals move from one type of household to another, and the households must be maintained in accordance with those moves.

While this modeling strategy makes sense, it implies an additional level of accounting (e.g. programming for the microsimulation) to maintain the households. I have instead chosen to directly model at the household level. Households will be defined in terms of the head of the household. The head is the man in a couple, the single parent, or the single individual living alone. Therefore when a household splits into two, the portion containing the original household head is the continuation of the original household. The other portion is considered a new household. They are not dropped from the analysis. If they were dropped or excluded, the estimation results would clearly be biased. Duncan and Hill's modeling strategy is a natural extension to the results in this paper, and I plan to estimate such a model in the future and compare the final simulation results between both techniques.

4 Introduction to Hazard Models

Hazard (or duration) models are used to model the time until an event or transition occurs. Most of the work in this field has been done by researchers in medicine and industrial engineering. Typical applications include models of drug effectiveness where the event of interest is curing the disease or death, and studies of machine reliability. Economists have applied the techniques to the study of unemployment duration (Meyer, 1990; Lancaster, 1979; Flinn and Heckman, 1982). More recently, duration models have been used in transportation analysis to study the timing of automobile purchases (Hensher and Mannering, 1994; Hensher, 1994; Jong, 1993) and accidents (Chang and Jovanis, 1990). They have also been applied to demographic processes (Heaton and Jacobson, 1994). The literature contains quite a bit of terminology, and the same concepts are often referred to differently in each field. As a point of reference, I will define the basic concepts as referred to in this paper.

A "state" describes the household's (or individual's) current status. For example, states might be single, married with 1 child, employed, unemployed, and so forth. The movement from one state to another is called a "transition" (or exit). The terms "spell" and "episode" are used to mean the total amount of time spent in a specific state before a transition occurs. If the household never makes a transition during the period of observation, the spell is said to be "right-censored". If the household was in the current state before the period of observation (and it is unknown when they entered their current state), then the spell is "left-censored".

The simplest hazard models describe a situation of only two states and one episode. A transition from the first state to the second only occurs once. The second state is referred to as an absorbing state such as death. More complex models include multiple states and multiple spells. Competing risk hazard models describe the situation where an spell can end in many different ways.

In this paper I will model demographic transitions at the household level using the nine possible states shown in Table 6. Roommates are not considered a household type: instead, each person would be considered a single adult.

Table 6: Household States

S	Single adult
C	Couple (including cohabiting adults together at least 1 year)
C1	Couple with 1 child
C2	Couple with 2 children
C3+	Couple with 3 or more children
S1	Single adult with 1 child
S2	Single adult with 2 children
S3+	Single adult with 3 or more children
O	Other households (extended families and first-year cohabiting adults)

The models will be estimated using the Panel Study of Income Dynamics (PSID), and I have used that data as a guide for determining how to break down the household types. Just under 5% of the PSID sample exists in the 'Other' state in any given year. Typical households described by this state include parents living with their children's family, adult brothers and sisters (and sometimes their children) living together, and first-year cohabiting adults. The following table shows the breakdown of other households. The percentages for 'Single' are for households that would have been coded as 'Single' or 'Single with children' if the additional household members were excluded. Likewise, the percentages for 'Couple' are for households that would have been coded as 'Couple' or 'Couple with children' if the additional household members were excluded.

Table 7: Breakdown of 'Other' Household Types
(Proportion by type if additional household members were excluded)

Additional household members:	'Single'	'Couple'	All Types
Brother or Sister	19.7	18.1	19.1
Parents	21.6	37.1	27.6
Spouse of child	5.8	2.4	4.5
First-year cohabitator	21.7	0	13.2
Other relative	11.4	30.1	18.7
Other non-relative	19.7	12.3	16.8

Based on PSID data from 1979 to 1980 broken down into 6-month intervals. 'Single' refers to households that would have been coded as single or single with children if the additional household members had been excluded. 'Couple' refers to households that would have been coded as couple or couple with children if the additional household members had been excluded.

Although additional states could be added for each such subtype, I have combined them all because there are few households of each subtype. Including further breakdowns could be a future extension to this work.

The PSID makes the following distinction for cohabiting adults: after one year, unmarried couples are treated as though they were married. According to Bumpass (1990), the average length of cohabitation is relatively short (the median is 1.5 years), and almost 50% of first marriages in the mid-1980's were preceded by a cohabitation period. It seems plausible that after one year most cohabiting couples are in a stable, "marriage-like" relationship, or at least well on their way to being legally married. Therefore, I have maintained the PSID convention of separating cohabiting adults, and including longer-term relationships (over 1 year) in the 'Couple' category. Recent cohabitators are included in the 'Other' category.

Additional states could also be added for larger families. The grouping of household with 3 or more children could be split into finer categories (e.g. 3 children, 4 children, 5+ children). Since there are few households in each finer categories, I have chosen to group them together.

Death can be included in several ways. 'Single' households may die by dropping them from the sample of households. Simple death tables tabulated by age can be used to simulate this event. In addition, transitions from 'couple' to 'single' may also be caused by death of the spouse. Simple death tables again can be used to determine whether to create the splitoff household caused by the divorce, or to assume that the spouse died.

Given these nine categories, households move from one type to another over the lifetime of household members. These movements will modeled with hazard models.

4.1 Basic Two-State Hazard Models

Before describing the more complex competing risk models, I will begin with some of the basic concepts for a simple single-episode hazard model with only two states.[6] T is defined as a continuous non-negative random variable. T = 0 when the household enters a given state, and represents the amount of time spent in that state (also referred to as the duration of stay).

The probability that the household exits the current state during some small time period (dt) given that it has been in current state for an amount of time equal to t is represented by $P(t \leq \mathrm{T} < t + dt \mid \mathrm{T} \geq t)$. Dividing this probability by dt gives the average probability of leaving per unit of time conditioned upon remaining in the state until t. As dt goes to zero, we are left

[6] Most of this section is based on Lancaster (1990). Chapter 2. Sections 2.1-2.3.

with the hazard function, *h(t)* given in 1.1, which represents the instantaneous rate of leaving at t conditioned upon remaining in the state until t.

$$h(t) = \lim_{dt \to 0} \frac{P(t \le T \le t + dt \mid T \ge t)}{dt} \tag{1.1}$$

The hazard function can also be expressed in terms of the distribution and density functions of T, $F(t)$ and $f(t)$ respectively. $F(t)$ is defined as P(T < t) as opposed to the usual P(T ≤ t). In addition, $F(0) = 0$ and $f(t) = \frac{d}{dt}F(t)$. This representation of the hazard function is shown in equation 1.2.[7]

$$h(t) = \frac{f(t)}{1 - F(t)} \tag{1.2}$$

Again, the hazard function is defined as the conditional density of the transition out of a state given the length of time already spent in that state.

The survivor function is defined as the probability that the household remains in the current state at least until time t:

$$S(t) = 1 - F(t) = \frac{f(t)}{h(t)} \tag{2}$$

Based on these equations, a given F() determines the hazard and survival functions. Additionally if we know the hazard function, we can determine the distribution and density

[7] Equation 1.2 can be derived by using the law conditional probability. $P(t \le T \le t + dt \mid T \ge t) =$ $P(t \le T \le t + dt, T \ge t)/P(T \ge t)$ which equals $P(t \le T \le t + dt)/P(T \ge t)$ since $T \ge t$ is a subset of $(t \le T \le t + dt)$. Using the distribution and density functions, $P(t \le T \le t + dt)/P(T \ge t) = F(t + dt) - F(t)/1 - F(t)$. Finally dividing by dt as dt goes to zero gives $h(t) = \lim_{dt \to 0} (F(t + dt) - F(t))/(1 - F(t))\, dt = F'(t)/(1 - F(t))$. The hazard function, $h(t) = f(t)/(1 - F(t))$.

functions for the failure time since the distribution function is simply the solution to the differential equation in 1.1.[8]

$$F(t) = 1 - \exp\left(-\int_0^t h(u)\,du\right) \tag{3}$$
$$f(t) = h(t)\exp\left(-\int_0^t h(u)\,du\right)$$

What does the hazard function look like and how should it be specified? There are two approaches, namely parametric and nonparametric specification. I will focus on the parametric specification. The appropriate functional form varies across applications, and is related to the concept of duration dependence. Duration dependence describes how the failure rate changes (or remains the same) as time passes. In the simplest case, the failure rate is constant implying that the probability of exit is independent of the length of the duration. Typically this case is modeled by defining f(t) as an exponential distribution so that $f(t) = \lambda\exp(-\lambda t)$ and $h(t) = \lambda$. In practical applications, a constant hazard rate is often too limiting. Instead positive or negative duration dependence may be more appropriate. Positive duration dependence exists when households are more likely to exit their current state as time passes, $\partial h/\partial t > 0$. Negative duration dependence exists when households are less likely to exit their current state as time passes, $\partial h/\partial t < 0$. Several functional forms for the hazard function capture either positive or negative duration dependence such as the Weibull or Gompertz models.

---

[8] The solution to the differential equation in 1.1 is derived as follows. First note that $f(t) = -\frac{d}{dt}[1-F(t)]$. The hazard function can be rewritten as $h(t) = -\frac{d}{dt}[1-F(t)]\cdot\frac{1}{1-F(t)}$. Integrating both sides gives $\int_0^t h(u)du = -\int_0^t d[1-F(u)]\cdot\frac{1}{1-F(u)}$ which evaluates to $\int_0^t h(u)du = -\ln[1-F(u)]\Big|_0^t = -\ln[1-F(t)]$ since F(0)=0. Rearranging leads to the first line of equation 3, $F(t) = 1-\exp(-\int_0^t h(u)du)$.

Davies (1992) shows that the duration dependence changes for many demographic transitions. For example, a single person might have positive duration dependence in their twenties and thirties, constant duration dependence for a period, and then negative duration dependence (e.g. past a certain point, if they are single, they will most likely never marry or remarry). In the early twenties, the person might be postponing marriage because of educational considerations but as they reach late twenties and early thirties they might want to marry in order to have children. As the "biological clock" ticks away, women experience positive duration dependence. Past a certain point, they are no longer marrying to have children, at which point constant duration dependence might kick in. Finally, the pool of available mates becomes smaller and smaller so that she may experience negative duration dependence as she ages. This reasoning is meant only to be suggestive. In such cases, the hazard function can be modeled as following a quadratic or cubic form.[9]

Flinn and Heckman (1982) suggest the following flexible parametric function form for the hazard function. The exponential form guarantees that the hazard function is nonnegative.

$$h(t) = \exp(\gamma_0 + \sum_{k=1}^{K} \gamma_k \left[\frac{t^{\lambda_k} - 1}{\lambda_k} \right]) \tag{4}$$

The term within brackets is a Box-Cox transformation of time. As λ_k approaches 0, this transformation approaches ln(t). The values for K and λ_k determine the specific functional form as given in Table 8. Constant, increasing, decreasing and varying duration dependence are all included in the flexible form depending upon the specific parameters values for K, λ_ks, and γ_ks.

[9] The software package (CTM) which I have used allows for exponential, Weibull, Gompertz, quadratic and cubic forms of the hazard function. Other potential options include log-logistic and log-normal forms. These are not currently available in CTM.

Table 8: Hazard Functions from the Flexible Parametric Form

Exponential (K=0)	$h(t) = \delta, \quad \delta = \exp(\gamma_0)$	(4.1)
Weibull (K=1, λ_1=0)	$h(t) = \delta t^{\gamma_1}, \quad \delta = \exp(\gamma_0)$	(4.2)
Gompertz (K=1, λ_1=1)	$h(t) = \delta \exp(\gamma_1 t), \quad \delta = \exp(\gamma_0 - \gamma_1)$	(4.3)
Quadratic (K=2, λ_1=1, λ_2=2)	$h(t) = \delta \exp(\gamma_1 t + \frac{\gamma_2}{2} t^2), \quad \delta = \exp(\gamma_0 - \gamma_1 - \frac{\gamma_2}{2})$	(4.4)
Cubic (K=3, λ_1=1 λ_2=2 λ_3=3)	$h(t) = \delta \exp(\gamma_1 t + \frac{\gamma_2}{2} t^2 + \frac{\gamma_3}{3} t^3), \quad \delta = \exp(\gamma_0 - \gamma_1 - \frac{\gamma_2}{2} - \frac{\gamma_3}{3})$	(4.5)

Household characteristics such as income, ethnicity, and age of the household members are likely to influence the rate of transitions. These characteristics are referred to as covariates, of which there are four major types. The first type do not vary over time, such as race, gender, and indicators of previous demographic status. These are referred to as time-invariant. The second type may vary over time, but their complete path is known before hand. For example, age is deterministic once the age at entry to the current state is known. These are referred to as "defined external covariates" (Kalbfleisch and Prentice, 1980) and can also be treated as time-invariant. The third type varies over time, but has an unknown future path. If we can determine the value of the covariate based on past history of the covariate, and the covariate can be described by a process that is unrelated to the duration of stay in current state, then it is referred to as an "exogenous" covariate (Lancaster, 1990). All of these three types can easily be incorporated into the hazard function as shown in equation 5. I assume that the covariates in this study are of these three types.

$$h(t,x) = \exp(\gamma_0 + x_t\beta + \sum_{k=1}^{K} \gamma_k \left[\frac{t^{\lambda_k} - 1}{\lambda_k}\right]) \tag{5}$$

x_t is a vector of exogenous covariates and β is an additional vector of parameters to be estimated. All of the five specific functional forms in Table 8 can be expanded to include covariates by replacing γ_0 by $\gamma_0 + x_t\beta$.

The final covariate type is referred to as "endogenous" (Lancaster, 1990). In this case, the fact that the current state has not been left by time $t + dt$ helps predict the covariate value from time t to $t + dt$. Endogenous covariates complicate the models, and "raise some rather subtle issues not all of which have been fully clarified in the literature" (Lancaster, 1990, page 23). I will assume that all covariate used in the models are not endogenous.

The previous models assume that the hazard function and survival distribution are homogeneous over the population of households. This assumption will almost certainly be invalid when important explanatory variables are excluded from the model, or when the transition times or covariates are imprecisely measured. In either case, the problem can be corrected by including an unobserved heterogeneity term in the hazard function.

Given these basic concepts, the demographic application that I am interested in requires a more complex model. Households can transition among a number of states, not just two, and multiple episodes are observed over the lifetime. A typical household may begin as a single person, transition into a couple, then to a couple with a child or possibly several children, and finally end as a couple. The variations are numerous. The most common transitions will be modeled using competing risk hazard models.

4.2 Competing Risk Hazard Models

4.2.1 Multistate/Single Episode

Again I begin with the simpler case of single-episode models.[10] The hazard function and the distribution and density functions of T are still defined as before. But now, the household may

[10] Most of this section is based on Lancaster, 1990, Chapter 5 and Flinn and Heckman, 1992.

leave the current state to several possible destinations which leads to conditional analogs to the hazard, distribution and density functions.

Assuming that the household begins in state i (i=1,2,..N), it can then move to one of N - 1 other states represented by j. Let D_j be a dummy variable that indicates whether state j was entered upon transition (e.g. D_j =1 if state j was entered, D_j = 0 otherwise). The "transition intensity" (sometimes referred to as a state-specific hazard function) represents the instantaneous rate of leaving state i to state j at time t conditioned upon remaining in state i until t. It is given by the following equation:

$$h_{ij}(t) = \lim_{dt \to 0} \frac{P(t \leq T \leq t + dt, D_j = 1 \mid T \geq t)}{dt} \tag{6}$$

The usual hazard function represents the instantaneous rate of leaving state i to any given state (conditioned upon remaining in state i until t), and is simply the sum of the state-specific transition intensities.

$$h_i(t) = \sum_{j=1, j \neq i}^{N} h_{ij}(t) \tag{7}$$

Another important concept is the marginal probability of a destination; in other words, the probability that when the household exits the current state, they move to state j. This marginal probability is represented by π_{ij}. Recall that the survival function S(t) = [1 - F(t)]. Therefore the marginal probability of a destination is given by:[11]

$$\pi_{ij} = \int_0^{\infty} S(u) h_{ij}(u) du \tag{8}$$

The sum of these marginal probabilities over all N-1 possible destination states equal 1, $\sum_{j=1, j \neq i}^{N} \pi_{ij} = 1$.

[11] This can be derived by noting that S(t) $h_{ij}(t)dt$ = P(survival to t) x P(exit to state j in the interval t + dt | survial to t). Integrating over all t gives the empirical counterpart of the fraction of households that ever transition to state j.

Finally, the conditional distribution of T (e.g. conditioned on transitioning to state j) given the starting state is i is represented by $F_{ij}(t)$. In other words, $F_{ij}(t)$ represents the probability that the household departs state i before time t given that when the departure occurs, it is to state j. Therefore $\pi_{ij} F_{ij}(t)$ is the probability that the household departs before time t *and* that they depart to state j. Finally the original unconditional distribution of T for the starting state i is given by the sum of this product over all possible j, $F_i(t) = \sum_{j=1, j \neq i}^{N} \pi_{ij} F_{ij}(t)$.

We only observe that the household left the initial state at some time t, and entered one of the j possible states. Using the dummy variables (D_j) and T, the likelihood function is given as follows:

$$\exp\left(-\int_0^t h_{ij}(u)du\right) \prod_{j=1, j \neq i}^{N} h_{ij}(t)^{D_j} \tag{9}$$

Another way of formulating this same model uses latent exit times from state i to the other N-1 possible states. This is what the terminology "competing risks" refers to. By assuming that these latent exit times are independent, the joint density of the those latent exit times is also given by equation 9 above. This assumption makes the estimation procedure much easier, but at may be unrealistic for some demographic transitions. For example, the transition from the state of couple with one child to the state of couple with two children may be related to the hazard rate for moving from the state of couple with one child to single with one child. In other words, the failure time for the first transition may be related to the latent failure time for second type of transition. Knowing that you are unlikely to divorce might influence your decision to have another child, just as knowing that you are on the verge of divorce affects your decision to have another child (Lillard, 1993). Unobserved heterogeneity may be used to address this issue by allowing for interdependencies among groups of transitions as described below.

Given these basic concepts and definitions, the flexible parametric form for the hazard function in equation 5 can be extended to deal with the multistate case.

$$h_{ij}(t,x) = \exp(\gamma_0 + x_t\beta_{ij} + \sum_{k=1}^{K} \gamma_{kij}\left[\frac{t^{\lambda_{kij}}-1}{\lambda_{kij}}\right]) \qquad (10)$$

This hazard function includes time-varying covariates and terms for duration dependence. An unobserved heterogeneity component, V_{ij}, can also be added. Flinn and Heckman (1982) suggest adding it to the term within the exponent.

$$h_{ij}(t,x) = \exp(\gamma_0 + x_t\beta_{ij} + \sum_{k=1}^{K} \gamma_{kij}\left[\frac{t^{\lambda_{kij}}-1}{\lambda_{kij}}\right] + V_{ij}) \qquad (11)$$

Unobserved heterogeneity is important when important explanatory variables have been excluded from the model. In this situation, groups of transitions are inherently inter-related, and the unobserved heterogeneity captures the interdependence.

Unobserved heterogeneity term can be specified in several ways. For ease of estimation, Flinn and Heckman (1982) suggest simplifying the unobserved heterogeneity components by assuming that they are constant within spells but vary across spells, so that unobserved heterogeneity across spells is restricted to a one-factor error specification.

$$V_{ij} = C_{ij}\, V \qquad (12)$$

The software which I have used to estimate the models (CTM) uses this form, and estimates the parameter C_{ij}. Some assumption must be made about the distribution of V. For example, V may be normally distributed with mean 0 and variance of 1, V~N(0,1). Therefore the variance of V_{ij} is allowed to change across spells. CTM allows several other assumptions about the unobserved term. V may be follow a lognormal, exponential, gamma, or even nonparametric distribution.

Unobserved heterogeneity may be used to the fact that the latent exit times are not truly independent. The independent competing risk model with unobserved heterogeneity can be estimated on subsets of the states (e.g. those states that are thought to be more closely related). A significant unobserved heterogeneity parameter (C_{ij}) implies that important variables have been excluded, and the transitions are inter-related.

The fullest hazard model is the multistate and multi-episode formulation. Multi-episodic data allows for variation in hazard functions depending on which episode is being modeled. For example, a second divorce may be different from the first.

4.2.2 Multistate/Multi-episode

Equations 6 through 12 can be expanded to a multi-episode model. Over the period of study, the household experiences several transitions across the possible states. Assume that the household is in state i for its m^{th} spell. It has been in this state for a length of time equal to t_m and entered the state at calendar time τ_m. The hazard function is given by:

$$h_{i_m j_m}(t_m, x) = \exp(\gamma_0 + x_{t_m + \tau_m}\beta_{ij}^m + \sum_{k=1}^{K} \gamma_{kij}\left[\frac{t_m^{\lambda_{kij}} - 1}{\lambda_{kij}}\right] + V_{ij}) \tag{13}$$

where the coefficient vector β_{ij}^m can vary across episodes.

The unobserved heterogeneity can again be simplified by assuming that $V_{ij} = C_{ij}V$ where several distributional assumptions can be made about V such as $V \sim N(0,1)$. The derivation of the likelihood function for this model can be found in Flinn and Heckman (1982, 1983).

5 The Panel Study of Income Dynamics

As was mentioned previously, the models will be estimated using data from the Panel Study of Income Dynamics (or PSID). The PSID began in 1968, and has surveyed the same sample of households every year. When children move out or families split apart, every effort was made to track both the original and the new splitoff household. Response rates are quite high, and range from 97 to 98.5% each year.

The PSID was initially conducted to study poverty. The full data set contains a large over sampling of low-income households, and can be broken into two subsamples: (1) the Social Research Council (SRC) sample of approximately 3000 household which was randomly drawn from the population of the 48 contiguous states, and (2) the non-random SEO sample of approximately 2000 low-income households selected from respondents to the Survey of Economic Opportunity. For the purposes of this demographic modeling, I have only used the random SRC sample for the years 1980 through 1989. The SRC subsample contained between 3500 and 4000 households each year. New households enter the sample as they split off from existing households, and others leave the sample because of nonresponse. Overall, the sample contains information for approximately 4600 households. In some cases (such as splitoff households), the data is available for only a subset of the full ten years.

The PSID collects information about movements into and out of households each month. I have set the observation period at every 6 months, based on the assumption that only 1 transition is likely to occur in that time frame. In other words, I have determined household type for each household in the PSID every 6 months. For the 4600 households, almost 10,000 spells occurred in the 10 year period. Table 9 shows a 9x9 transition matrix illustrating all the possible transitions and frequency counts. The headings on the left side of each row indicate the beginning state, and the headings across the top of the columns indicate the ending state. Highlighted cells indicate transitions that were estimated. For example, the transition from 'Couple with 2 children' to 'Single' is included, but the transition from 'Single' to 'Couple with 2 children' is not. I do not plan to estimate several cells, particularly those with less than 50 observations. Transitions with over 50 observations which currently are excluded may be included in future estimations (mainly household types described as 'Other').

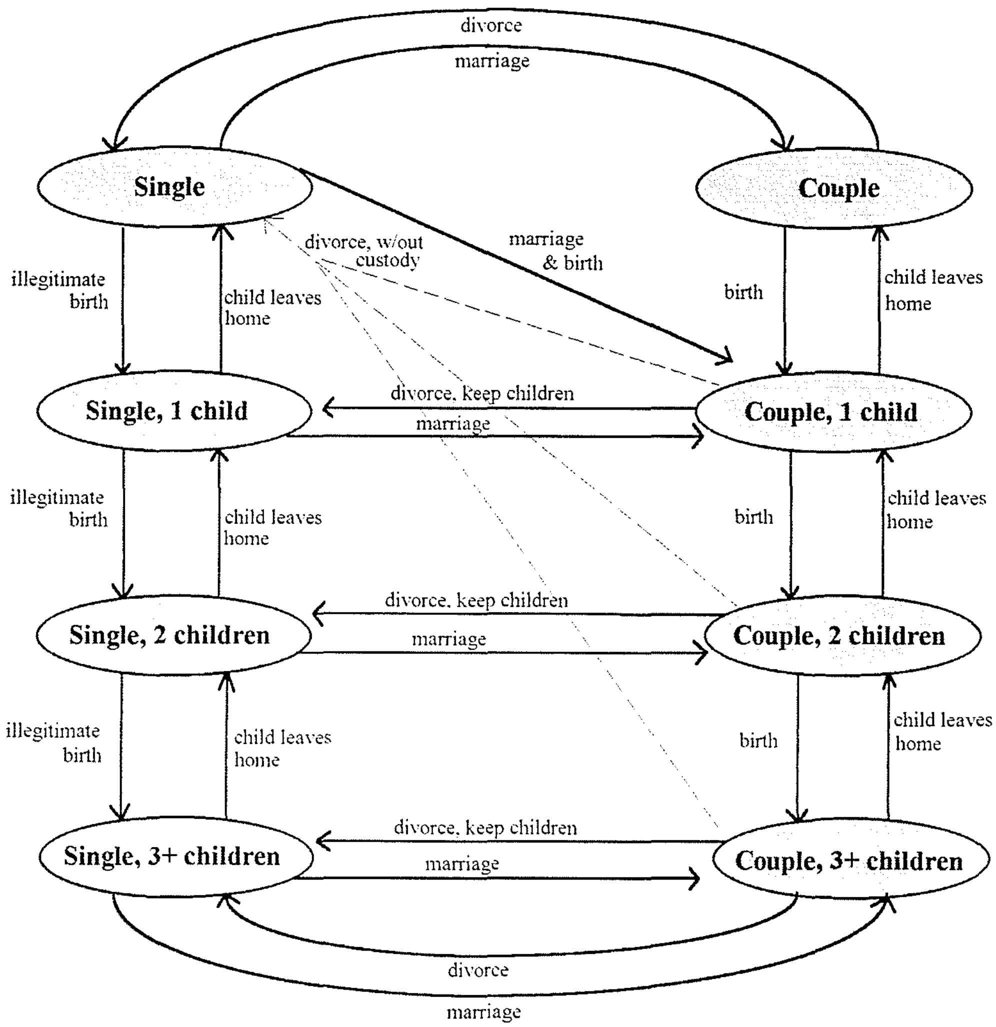
Table 9: Household Transition Matrix and Frequency Counts

States	S	C	C1	C2	C3+	S1	S2	S3+	O
S	1192	428	67	31	15	119	8	4	187
C	363	1063	634	17	5	5	0	0	59
C1	78	311	591	671	17	115	8	0	49
C2	64	47	303	674	295	18	89	3	50
C3+	23	6	31	187	412	8	4	52	39
S1	119	12	95	13	8	185	55	0	35
S2	18	2	9	52	9	64	110	34	22
S3+	4	3	3	5	24	9	44	59	9
O	156	116	54	46	38	23	16	10	193

It is important to note that just because a transition is not directly modeled, households may still make those moves, just not in a 6 month time frame. For example, a 'Single' household may move to 'Couple with 2 children' by first transitioning to 'Couple with 1 child' and then moving from 'Couple with 1 child' to 'Couple with 2 children'. At a minimum, this path would take 1 year.

Figure 1 shows the possible movements among all household states except 'Other'. Movements labeled "marriage" also include long-term cohabitation arrangements. Couples with children can dissolve in two ways, either the household head has custody of the children (the solid lines) or the he leaves the relationship without custody (the dashed lines). In the first case, the remaining adult forms a new single household. In the second case, the remaining adult forms a new household consisting of a single with children.

Figure 1: Flow Chart of Demographic Changes



The PSID is rich with possible explanatory variables, many of which have been identified in the demographic literature as being significant. But I must be able to forecast all explanatory variables. Therefore I have limited the covariates to race, education level, employment status, age and gender of adult household members, age of children, and household income. For transitions beginning from a 'couple' state, I have included information about both the husband and wife.

5.1 Race

In 1980, the PSID random sample contains a breakdown of 88.4% white households, 9.4% black households, and 2.2% of other households (mainly Asian, with some Native American). In 1989 the breakdown has shifted slightly to include 90.3% white households, 8.4% black households, and 1.3% other races. Since blacks are more likely to postpone marriage, bear children out-of-wedlock, and experience martial disruption, and live in extended families, I expect that a large percentage of black households would be classified as 'single with children' or 'other'. Table 10 shows racial breakdown for each household type for the years 1980 and 1988.

Table 10: Race by Household Type

1980	All	C	C1	C2	C3+	S	S1	S2	S3+	O
% White	88.4	93.9	87.9	91.4	86.6	90.1	74.1	73.3	57.1	74.5
% Black	9.4	5.3	9.2	5.5	10.8	8.6	23.0	22.1	40.8	19.8
% Other	2.2	0.8	2.9	3.1	2.6	1.3	2.9	4.6	2.1	5.7
Total households	3579	900	555	514	344	839	135	86	49	157
1989										
% White	90.3	94.4	93.5	93.9	88.7	90.2	76.2	68.5	44.2	87.0
% Black	8.4	4.6	4.8	5.1	9.4	8.6	23.1	29.2	51.9	11.7
% Other	1.3	1.0	1.7	1.0	1.9	1.2	0.7	2.3	3.9	1.3
Total households	3802	945	525	607	371	904	147	89	52	162

S - Single, C - Couple, C1 - Couple with 1 child, C2 - Couple with 2 children, C3 +- Couple with 3 or more children, S1 - Single with 1 child, S2 - Single with 2 children, S3+ - Single with 3 or more children, O - Other household types

As expected, more blacks are categorized as 'single with 1 child', 'single with 2 children', and 'single with 3+ children' than the black proportion of the population as a whole. In the most extreme case, 50% of all PSID households consisting of single of 3 or more children are headed by a black person in 1989. Only 44.2% of such households are headed by a white person in that same year while whites make up 90.3% of the total sample.

Other the other hand, whites make up a larger relative percentage of couples, couples with 1 child, and couples with 2 children. For example, 94.4% of couples are headed by a white person in 1989 while whites make up 90.3% of the total sample in that year.

5.2 Education

The demographic literature also provided insight into the relationship between education and household transitions. Those with less education are more likely to divorce (with the exception of women with advanced graduate work), are more likely to remarry for women and less likely to remarry for men, and are less likely to delay child birth. Table 11 shows the breakdown of education for each household type and for the sample as a whole.

Overall, education levels were rising from 1980 to 1989. In 1980 very few people had continued past college. By 1989, 10.6% had continued on to graduate work. Likewise the percentages for those without a high school diploma had fallen sharply from 25.3% in 1980 to 16.2% in 1989.

Table 11: Education by Household Type

1980	All	C Husb	C Wife	C1 Husb	C1 Wife	C2 Husb	C2 Wife	C3+ Husb	C3+ Wife	S	S1	S2	S3+	O
% Some HS	25.3	30.3	23.9	19.4	17.2	20.0	19.3	23.5	24.8	31.4	41.5	41.2	58.3	32.1
% HS grad	40.6	33.3	43.8	41.7	54.5	37.8	51.4	36.7	51.6	31.1	37.8	38.8	31.3	34.6
% College	33.4	35.8	31.6	38.1	29.8	41.7	28.8	39.3	22.7	36.8	20.7	20.0	10.4	33.3
% Graduate	0.7	0.6	0.7	0.8	1.5	0.5	0.5	0.5	0.9	0.7	0.0	0.0	0.0	0.0
1989														
% Some HS	16.2	19.3	15.3	15.2	12.0	10.1	10.1	12.5	11.3	24.1	20.4	28.1	37.3	23.5
% HS grad	38.8	34.9	40.0	38.1	47.6	36.1	44.1	40.1	46.1	31.3	44.2	36.0	33.3	37.7
% College	34.4	31.5	34.9	37.9	32.6	40.4	35.6	35.2	35.0	33.9	28.6	25.8	25.5	28.4
% Graduate	10.6	14.3	9.8	8.8	7.8	13.4	10.2	12.2	7.6	10.8	6.8	10.1	3.9	10.5

S - Single, C - Couple, C1 - Couple with 1 child, C2 - Couple with 2 children, C3 +- Couple with 3 or more children, S1 - Single with 1 child, S2 - Single with 2 children, S3+ - Single with 3 or more children, O - Other household types. For couple household types, ages are shown for both husbands (husb) and wifes.

For all of the single with children categories, the percentage of household heads with low education levels is higher than for the population as a whole. For example, 37.3% of households which are single with 3 or more children are headed by a person who has not graduated from high school. The corresponding percentage for the population as whole is 16.2%.

Men with a higher level of education (college or graduate work) are more likely to be married with children. In 1989 37.9% of couples with 1 child, 40.4% of couples with 2 children, and 35.2% of couples with 3 or more children contain a husband with some college education. These percentages are higher than the population as a whole which contains 34.4% of people with some college education.

In general, men within couples appear to have more schooling than their wives at the higher categories of education. On the other hand, women in couples are more likely to have finished high school than their husbands at the lower categories of education.

5.3 Other Covariates

I have included several other covariates besides race and education. These covariates include household income, employment status, age, gender of singles, number of children in various age categories, and some indicators of previous states occupied. Covariates were included only when appropriate (e.g. number of children was not included for transitions out of the 'single', 'couple', or 'other' states).

Household income for the years 1979 through 1989 was converted to 1989 dollars, and the income for the entire year was divided by 2 to represent 6-month earnings. But using income for the current year that the transition might have occurred would lead to biased results. For example, a single person who marries would necessarily have higher income during that year if his or her spouse worked. So instead of using income from the current year, I have used lagged income from the prior year. Employment status was coded as either employed, unemployed, or out of the workforce (which includes homemakers). The age categories for children included the number of children less than 6 years old, number of children between 6 and 18 years old, and number of children over 18 years.

Multi-episode modeling was used by included indicators of previous states occupied. A flag for previously married was set if the household was in any of the couple states during some past period of observation. Another flag was set if the household previously existed in a state of single with children. I did not use data or information prior to 1979 in setting these flags.

6 Estimation Results

6.1 Without Unobserved Heterogeneity

I have estimated models for each highlighted transition in Table 9 assuming independent Weibull hazard functions without unobserved heterogeneity. The results are summarized in the following tables. Tables 12.1 though 12.5 include transitions which begin in either the 'single' or 'other' states (e.g. states S, S1,S2,S3+, and O). Tables 12.6 though 12.9 show the results for transitions which begin in 'couple' states (e.g. C, C1, C2, and C3+). All tables include general covariates describing the household such as income, while tables 12.1-12.5 include covariates describing the single household head and tables 12.6-12.9 include covariates describing the both adult members in the household referred to as husband and wife.

For dummy variables, a negative coefficient implies that the group will remain in the current state longer than the reference group. In other words, it will take them longer to transition. Race is one such dummy variable. The coefficient for blacks in the model of transitions from 'Single' to 'Couple' is negative (-0.777 in Table 12.1 with a t-statistic of 2.79). This implies that blacks will remain in the single state longer than whites (the control group) before getting married. Blacks are likely to marry at a later age supporting the results of Heaton and Jacobson (1994).

A positive coefficient for a dummy variable implies that the group will transition faster than the reference group. Previous research finds that blacks are more likely to have a child out-of-wedlock, and at an earlier age than whites. My results support those findings. The coefficient for blacks in the the model of transititioning from 'Single' to 'Single with 1 child' is positive and significant (1.424 in Table 12.1 with a t-statistic of 6.52).

For continuous variables, a negative coefficient implies that as the variable increases, the household is likely to remain in the current state longer. Likewise, a positive coefficient implies that the household is likely to transition sooner at larger values of the variable. For example, at higher income levels (e.g. larger values for the log of income), singles are more likely to marry sooner (the coefficient is 0.196 in Table 12.1 with a significant t-statistic).

Table 12.1: Estimation Results for Independent Hazard Functions
Weibull Duration Dependence without Unobserved Heterogeneity

	Single → Couple		Single → Couple, 1 child		Single → Single, 1 child	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Constant intercept (γ_0)	-3.509	5.32	-6.857	3.03	-4.812	4.20
Gamma_1 (γ_1)	0.275	4.08	0.142	0.71	-0.087	0.73
Log (lagged household income)	0.196	3.29	0.079	0.54	0.125	0.94
Race (vs. white) [a]						
- Black	-0.777	2.79	-0.334	0.71	1.424	6.52
- Other	-0.286	1.68	-0.493	1.04	-0.146	0.40
Single's education (vs. < high school)						
- High school graduate	-0.054	0.28	-0.866	2.74	-0.261	1.05
- College	-0.019	0.10	-1.551	4.26	-0.755	2.81
- Graduate work or degree	0.111	0.48	-2.315	2.88	-1.303	2.56
Single's employment (vs. employed)						
- Unemployed	0.058	0.27	0.178	0.45	0.644	2.17
- Out of work force	0.166	0.78	-0.016	0.03	0.516	1.79
Age of single	-0.055	2.30	0.155	1.29	-0.030	0.96
(Age of single)^2	0.000	0.10	-0.003	1.51	0.000	0.23
Gender of single (female vs. male)	-0.200	2.01	-1.651	3.51	0.599	2.70
# of kids < 6 years old	----	----	----	----	----	----
# of kids between 6 & 18 years old	----	----	----	----	----	----
# of kids >= 18 years old	----	----	----	----	----	----
Previous marriage	0.369	3.16	0.926	3.51	----	----
Previously a single parent	----	----	----	----	----	----
- Log likelihood($\hat{\gamma}_0, \hat{\gamma}_1, \hat{\beta}$)	1770.81		359.07		636.61	
- Log likelihood($\hat{\gamma}_0, \hat{\gamma}_1$)	1974.82		431.68		696.71	

[a] Race is only reported for the head of the household. For single households, this poses no uncertainty. For couples, the head may be male or female. In --% of couples, the man has been coded as the head of the household.

Table 12.2: Estimation Results for Independent Hazard Functions
Weibull Duration Dependence without Unobserved Heterogeneity

	Single → Other		Single, 1 child → Single		Single, 1 child → Couple, 1 child	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Constant intercept (γ_0)	-4.919	6.54	-6.525	3.41	-0.240	0.13
Gamma_1 (γ_1)	0.007	0.08	0.165	1.32	0.259	1.59
Log (lagged household income)	0.181	2.94	0.028	0.17	0.065	0.63
Race (vs. white) [a]						
- Black	0.089	0.32	-0.106	0.41	-0.778	2.81
- Other	-0.161	0.63	-0.353	0.91	-0.053	0.15
Single's education (vs. < high school)						
- High school graduate	-0.264	1.15	-0.759	3.27	0.476	1.70
- College	-0.379	1.70	-0.741	2.90	0.680	1.96
- Graduate work or degree	-0.507	1.44	-0.740	1.69	0.887	1.72
Single's employment (vs. employed)						
- Unemployed	0.585	2.36	0.742	1.53	0.082	0.24
- Out of work force	0.377	1.38	-0.134	0.54	0.134	0.47
Age of single	-0.002	0.08	0.156	3.50	-0.047	0.45
(Age of single)^2	-0.000	1.37	-0.001	3.66	-0.001	0.86
Gender of single (female vs. male)	-0.172	1.15	-1.003	4.37	-1.201	4.08
# of kids < 6 years old	----	----	----	----	0.035	0.11
# of kids between 6 & 18 years old	----	----	----	----	----	----
# of kids >= 18 years old	----	----	0.828	2.92	----	----
Previous marriage	----	----	----	----	-0.200	1.01
Previously a single parent	----	----	----	----	----	----
- Log likelihood($\hat{\gamma}_0, \hat{\gamma}_1, \hat{\beta}$)	960.34		428.71		333.33	
- Log likelihood($\hat{\gamma}_0, \hat{\gamma}_1$)	1014.52		488.24		411.44	

[a] Race is only reported for the head of the household. For single households, this poses no uncertainty. For couples, the head may be male or female. In --% of couples, the man has been coded as the head of the household.

Table 12.3: Estimation Results for Independent Hazard Functions Weibull Duration Dependence without Unobserved Heterogeneity						
	Single, 1 child → Single, 2 children		Single, 2 children → Couple, 2 kids		Single, 2 children → Single, 1 child	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Constant intercept (γ_0)	-4.105	2.24	-7.152	1.64	-10.875	3.75
Gamma_1 (γ_1)	0.039	0.20	0.474	2.23	0.275	1.56
Log (lagged household income)	-0.139	1.09	0.487	1.97	0.162	0.73
Race (vs. white) [a]						
- Black	0.834	2.55	-1.486	2.17	-0.121	0.33
- Other	0.357	0.72	0.000	0.00	-0.307	0.68
Single's education (vs. < high school)						
- High school graduate	-0.094	0.25	0.247	0.71	0.286	0.70
- College	0.138	0.34	0.057	0.14	-0.211	0.44
- Graduate work or degree	-0.672	0.54	---	---	-0.654	0.86
Single's employment (vs. employed)						
- Unemployed	0.422	0.73	0.238	0.43	0.299	0.57
- Out of work force	0.556	1.54	0.477	1.27	-0.314	0.72
Age of single	0.077	1.27	0.063	0.26	0.225	2.54
(Age of single)^2	-0.001	1.25	-0.002	0.60	-0.002	2.54
Gender of single (female vs. male)	-0.506	1.12	-0.498	0.97	-0.134	0.31
# of kids < 6 years old	----	----	----	----	----	----
# of kids between 6 & 18 years old	----	----	-0.242	0.87	-0.178	0.39
# of kids >= 18 years old	-0.787	1.87	-0.647	0.99	0.697	1.40
Previous marriage	----	----	0.239	0.71	----	----
Previously a single parent	----	----	----	----	----	----
- Log likelihood($\hat{\gamma}_0, \hat{\gamma}_1, \hat{\beta}$)	256.27		202.30		231.97	
- Log likelihood($\hat{\gamma}_0, \hat{\gamma}_1$)	268.44		232.23		271.78	

[a] Race is only reported for the head of the household. For single households, this poses no uncertainty. For couples, the head may be male or female. In --% of couples, the man has been coded as the head of the household.

Table 12.4: Estimation Results for Independent Hazard Functions
Weibull Duration Dependence without Unobserved Heterogeneity

	Single, 2 children → Single, 3+ kids		Single, 3+ kids → Single, 2 children		Other → Single	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Constant intercept (γ_0)	0.256	0.09	-4.092	1.15	1.999	2.55
Gamma_1 (γ_1)	0.159	0.62	0.116	0.52	0.363	3.29
Log (lagged household income)	-0.278	1.72	-0.044	0.15	-0.130	1.76
Race (vs. white) [a]						
- Black	1.446	2.71	0.061	0.14	-0.317	-1.27
- Other	0.973	1.43	0.184	0.30	0.272	1.18
Single's education (vs. < high school)						
- High school graduate	-0.053	0.09	-1.001	2.10	0.486	2.03
- College	0.275	0.50	-0.833	1.68	0.830	3.59
- Graduate work or degree	----	----	----	----	0.713	1.95
Single's employment (vs. employed)						
- Unemployed	-0.587	0.71	0.310	0.46	0.381	1.21
- Out of work force	0.258	0.49	0.330	0.73	0.262	1.06
Age of single	-0.080	0.54	0.210	1.97	-0.177	6.86
(Age of single)^2	0.000	0.17	-0.002	0.15	0.001	5.19
Gender of single (female vs. male)	----	----	----	----	0.182	1.13
# of kids < 6 years old	----	----	-1.233	3.69	----	----
# of kids between 6 & 18 years old	-0.456	1.07	-1.062	3.57	----	----
# of kids >= 18 years old	0.347	0.67	-0.562	1.96	----	----
Previous marriage	----	----	----	----	----	----
Previously a single parent	----	----	----	----	----	----
- Log likelihood($\hat{\gamma}_0, \hat{\gamma}_1, \hat{\beta}$)	134.87		156.17		534.99	
- Log likelihood($\hat{\gamma}_0, \hat{\gamma}_1$)	166.76		180.26		606.20	

[a] Race is only reported for the head of the household. For single households, this poses no uncertainty. For couples, the head may be male or female. In --% of couples, the man has been coded as the head of the household.

Table 12.5: Estimation Results for Independent Hazard Functions
Weibull Duration Dependence without Unobserved Heterogeneity

	Other → Couple		Other → Couple, 1 child	
	Estimate	t-stat	Estimate	t-stat
Constant intercept (γ_0)	1.422	1.51	-1.501	0.67
Gamma_1 (γ_1)	0.023	0.15	0.065	0.29
Log (lagged household income)	-0.174	2.09	-0.004	0.02
Race (vs. white) [a]				
- Black	-1.337	2.88	-1.160	1.86
- Other	-0.165	0.59	-0.359	0.78
Single's education (vs. < high school)				
- High school graduate	0.405	1.54	-0.461	1.32
- College	0.423	1.52	-0.472	1.25
- Graduate work or degree	0.614	1.65	----	----
Single's employment (vs. employed)				
- Unemployed	0.026	0.07	-0.241	0.38
- Out of work force	-0.109	0.32	-0.612	1.00
Age of single	-0.113	3.17	-0.036	0.47
(Age of single)^2	0.001	2.31	-0.000	0.08
Gender of single (female vs. male)	-0.630	2.83	-0.686	1.74
# of kids < 6 years old	----	----	----	----
# of kids between 6 & 18 years old	----	----	----	----
# of kids >= 18 years old	----	----	----	----
Previous marriage	----	----	----	----
Previously a single parent	----	----	----	----
- Log likelihood($\hat{\gamma}_0$, $\hat{\gamma}_1$, $\hat{\beta}$)	442.22		247.86	
- Log likelihood($\hat{\gamma}_0$, $\hat{\gamma}_1$)	481.55		265.82	

[a] Race is only reported for the head of the household. For single households, this poses no uncertainty. For couples, the head may be male or female. In --% of couples, the man has been coded as the head of the household.

Table 12.6: Estimation Results for Independent Hazard Functions

Weibull Duration Dependence without Unobserved Heterogeneity

	Couple → Single		Couple → Couple, 1 child		Couple, 1 child → Single	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Constant intercept (γ_0)	0.799	1.03	-1.696	2.89	-0.776	0.31
Gamma_1 (γ_1)	0.274	3.47	0.334	6.24	0.374	1.97
Log (lagged household income)	-0.169	2.03	0.085	1.37	-0.068	0.34
Race (vs. white) [a]						
- Black	0.214	0.90	0.100	0.53	0.267	0.53
- Other	-0.250	1.26	-0.168	1.09	-0.648	1.26
Husband's education [b]						
- High school graduate	-0.203	1.27	-0.156	1.18	-0.575	1.69
- College	-0.323	1.83	0.037	0.26	-1.088	2.64
- Graduate work or degree	-0.236	1.11	-0.137	0.68	-0.987	1.46
Husband's employment						
- Unemployed	0.257	0.91	-0.009	0.05	1.499	4.19
- Out of work force	0.771	3.87	-0.781	3.53	1.090	2.11
Age of husband	-0.069	2.13	-0.099	2.95	0.026	0.23
(Age of husband)^2	0.001	1.87	0.001	1.62	-0.001	0.72
Wife's education						
- High school graduate	-0.105	0.59	0.048	0.34	-0.206	0.55
- College	0.069	0.34	0.052	0.32	-0.078	0.16
- Graduate work or degree	0.783	3.51	-0.005	0.02	1.442	2.63
Wife's employment						
- Unemployed	0.492	1.66	0.920	5.34	-0.558	0.72
- Out of work force	-0.058	0.36	1.386	15.34	-0.508	1.69
Age of wife	-0.080	2.32	0.015	0.39	-0.123	1.00
(Age of wife)^2	0.001	2.35	-0.001	1.37	0.001	0.82
# of kids < 6 years old	----	----	----	----	-0.571	1.49
# of kids between 6 & 18 years old	----	----	----	----	----	----
# of kids >= 18 years old	----	----	----	----	----	----
Previous marriage	----	----	----	----	----	----
Previously a single parent	----	----	----	----	0.155	0.39
- Log likelihood($\hat{\gamma}_0, \hat{\gamma}_1, \hat{\beta}$)	1518.48		2175.21		394.27	
- Log likelihood($\hat{\gamma}_0, \hat{\gamma}_1$)	1751.98		2708.82		458.44	

[a] Race is only reported for the head of the household. For single households, this poses no uncertainty. For couples, the head may be male or female. In --% of couples, the man has been coded as the head of the household.
[b] For all couples (even cohabiting adults). I refer to the male as the husband and the female as the wife.

Table 12.7: Estimation Results for Independent Hazard Functions
Weibull Duration Dependence without Unobserved Heterogeneity

	Couple, 1 child → Couple		Couple, 1 child → Couple, 2 children		Couple, 1 child → Single, 1 child	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Constant intercept (γ_0)	-10.798	8.33	-4.046	5.07	-0.747	0.43
Gamma_1 (γ_1)	0.176	2.44	0.539	9.62	0.330	2.28
Log (lagged household income)	-0.017	0.31	0.049	0.84	-0.138	0.78
Race (vs. white) [a]						
- Black	-0.003	0.01	0.172	1.07	0.565	1.67
- Other	0.120	0.63	-0.211	1.53	-0.622	1.36
Husband's education [b]						
- High school graduate	0.127	0.80	-0.042	0.31	0.192	0.70
- College	0.163	0.88	-0.030	0.21	-0.284	0.80
- Graduate work or degree	0.349	1.28	0.087	0.41	0.778	2.07
Husband's employment						
- Unemployed	0.305	0.89	0.161	0.87	0.751	1.82
- Out of work force	-0.217	1.05	-0.260	0.98	1.975	6.82
Age of husband	0.116	1.64	-0.084	1.98	-0.068	0.70
(Age of husband)^2	-0.001	1.24	0.001	1.58	-0.000	0.01
Wife's education						
- High school graduate	-0.277	1.77	-0.185	1.37	-0.187	0.66
- College	-0.039	0.19	-0.021	0.14	-0.098	0.28
- Graduate work or degree	-0.458	1.40	0.182	0.78	-0.154	0.29
Wife's employment						
- Unemployed	-0.246	0.52	0.687	3.66	-0.304	0.46
- Out of work force	0.005	0.39	0.657	7.76	-0.687	2.44
Age of wife	0.105	1.64	0.101	1.96	-0.056	0.60
(Age of wife)^2	-0.001	1.76	-0.002	2.51	0.001	0.98
# of kids < 6 years old	----	----	1.006	8.38	-0.035	0.10
# of kids between 6 & 18 years old	----	----	----	----	----	----
# of kids >= 18 years old	2.391	10.76	----	----	----	----
Previous marriage	----	----	----	----	----	----
Previously a single parent	----	----	----	----		
- Log likelihood($\hat{\gamma}_0, \hat{\gamma}_1, \hat{\beta}$)	1070.89		2203.47		533.49	
- Log likelihood($\hat{\gamma}_0, \hat{\gamma}_1$)	1384.40		2498.12		630.33	

[a] Race is only reported for the head of the household. For single households, this poses no uncertainty. For couples, the head may be male or female. In --% of couples, the man has been coded as the head of the household.
[b] For all couples (even cohabiting adults), I refer to the male as the husband and the female as the wife.

Table 12.8: Estimation Results for Independent Hazard Functions
Weibull Duration Dependence without Unobserved Heterogeneity

	Couple, 2 children → Single		Couple, 2 children → Couple, 1 child		Couple, 2 children → Couple, 3+ kids	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Constant intercept (γ_0)	3.439	1.11	-13.826	9.35	-0.368	0.27
Gamma_1 (γ_1)	0.308	1.42	0.066	0.91	0.197	2.26
Log (lagged household income)	-0.203	2.09	-0.035	0.49	-0.055	0.58
Race (vs. white) [a]						
- Black	0.270	0.42	-0.291	1.15	0.519	2.11
- Other	-0.357	0.62	0.125	0.65	0.145	0.79
Husband's education [b]						
- High school graduate	-0.146	0.30	-0.043	0.26	0.035	0.19
- College	0.413	0.82	-0.314	1.54	0.086	0.42
- Graduate work or degree	0.661	0.93	-0.268	0.92	-0.056	0.18
Husband's employment						
- Unemployed	0.467	0.95	0.698	2.60	0.201	0.82
- Out of work force	-0.690	0.44	-0.202	0.78	0.220	0.58
Age of husband	-0.046	0.23	0.138	1.94	-0.034	0.52
(Age of husband)^2	0.001	0.20	-0.001	1.46	0.000	0.40
Wife's education						
- High school graduate	0.145	0.26	-0.215	1.37	0.009	0.05
- College	0.022	0.33	-0.224	1.00	0.300	1.31
- Graduate work or degree	0.663	0.82	-0.070	0.21	0.304	0.90
Wife's employment						
- Unemployed	0.356	0.59	-0.029	0.06	0.702	2.21
- Out of work force	-1.075	2.54	0.055	0.41	0.981	7.30
Age of wife	-0.294	1.53	0.224	2.76	-0.144	2.01
(Age of wife)^2	0.003	1.17	-0.003	2.87	0.001	1.67
# of kids < 6 years old	-0.039	0.57	----	----	0.154	0.65
# of kids between 6 & 18 years old	0.025	0.05	0.799	2.83	0.050	0.24
# of kids >= 18 years old	----	----	1.803	6.25	----	----
Previous marriage	----	----	----	----	----	----
Previously a single parent	-0.029	0.05	----	----	----	----
- Log likelihood($\hat{\gamma}_0$, $\hat{\gamma}_1$, $\hat{\beta}$)	357.99		1070.40		1273.39	
- Log likelihood($\hat{\gamma}_0$, $\hat{\gamma}_1$)	393.18		1392.87		1366.12	

[a] Race is only reported for the head of the household. For single households, this poses no uncertainty. For couples, the head may be male or female. In --% of couples, the man has been coded as the head of the household.
[b] For all couples (even cohabiting adults), I refer to the male as the husband and the female as the wife.

Table 12.9: Estimation Results for Independent Hazard Functions
Weibull Duration Dependence without Unobserved Heterogeneity

	Couple, 2 children → Single, 2 kids		Couple 3+ kids → Couple, 2 children		Couple, 3+ kids → Single, 3+ kids	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Constant intercept (γ_0)	3.872	1.28	-7.970	2.74	1.510	0.37
Gamma_1 (γ_1)	0.438	2.60	0.057	0.57	0.319	1.15
Log (lagged household income)	-0.083	0.35	-0.080	0.93	0.053	0.20
Race (vs. white) [a]						
- Black	0.365	0.79	-0.171	0.45	0.692	1.33
- Other	-0.041	0.12	0.047	0.18	-0.075	0.15
Husband's education [b]						
- High school graduate	0.356	0.98	-0.345	1.58	-0.496	1.10
- College	0.303	0.75	-0.388	1.53	0.806	1.55
- Graduate work or degree	1.504	3.23	-1.133	2.99	0.273	0.33
Husband's employment						
- Unemployed	0.871	1.97	-0.041	0.11	0.868	1.54
- Out of work force	2.262	5.83	0.077	0.21	2.850	5.55
Age of husband	-0.037	0.30	0.216	2.11	0.366	1.99
(Age of husband)^2	0.000	0.22	-0.002	1.69	-0.004	2.02
Wife's education						
- High school graduate	-0.085	0.24	-0.122	0.49	0.030	0.07
- College	-0.434	0.96	-0.111	0.38	-0.404	0.59
- Graduate work or degree	-0.719	1.04	-0.337	0.62	0.241	0.27
Wife's employment						
- Unemployed	1.386	4.29	0.788	1.78	1.022	1.40
- Out of work force	-0.255	0.83	0.058	0.33	-0.227	0.56
Age of wife	-0.315	2.15	0.200	1.78	-0.707	3.46
(Age of wife)^2	0.003	1.81	-0.002	1.76	0.007	3.22
# of kids < 6 years old	-0.966	2.48	-1.720	4.86	-0.019	0.05
# of kids between 6 & 18 years old	-0.370	1.32	-1.572	5.39	-0.179	0.54
# of kids >= 18 years old	----	----	-0.960	3.46	0.156	0.41
Previous marriage	----	----	----	----	----	----
Previously a single parent	0.464	1.42	----	----	0.980	2.13
- Log likelihood($\hat{\gamma}_0, \hat{\gamma}_1, \hat{\beta}$)	431.35		674.47		238.70	
- Log likelihood($\hat{\gamma}_0, \hat{\gamma}_1$)	515.30		850.04		303.47	

[a] Race is only reported for the head of the household. For single households, this poses no uncertainty. For couples, the head may be male or female. In --% of couples, the man has been coded as the head of the household.
[b] For all couples (even cohabiting adults), I refer to the male as the husband and the female as the wife.

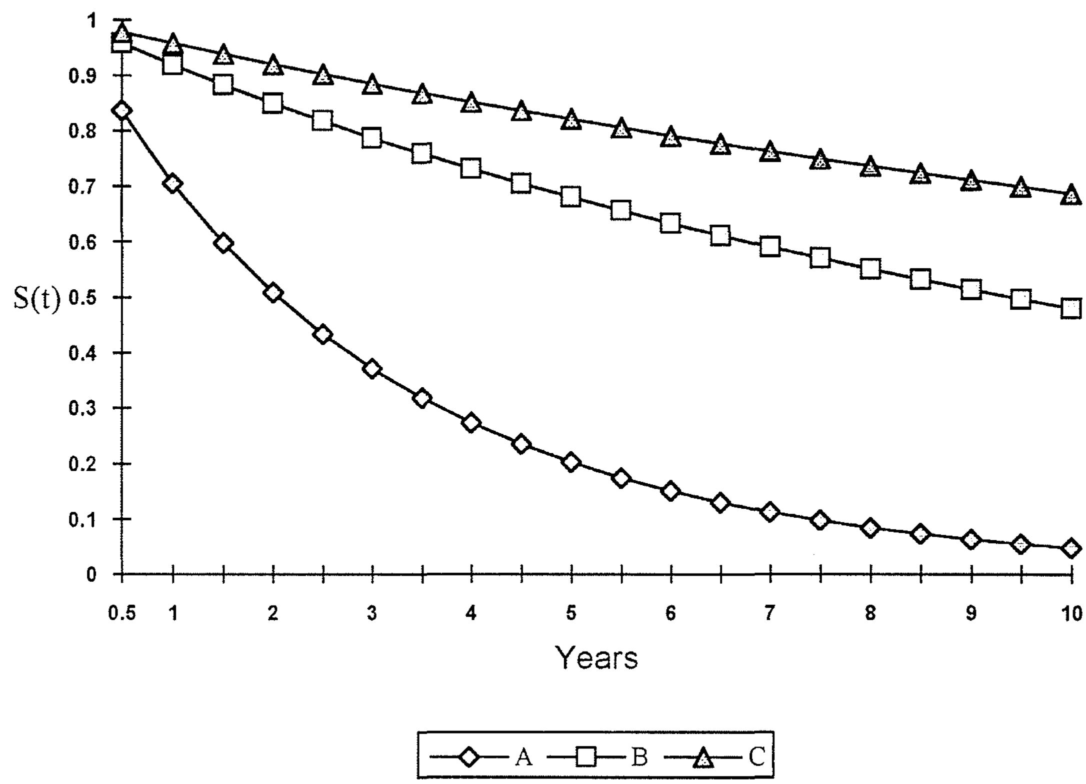
Are these models better than simple Weibull models without covariates? The likelihood ratio test answers this question by testing the null hypothesis that the coefficients on the covariates are zero, H_0: $\hat{\beta} = 0$. For all transitions except from 'Single with 1 child' to 'Single with 2 children', the likelihood ratio test rejects the null hypothesis that the coefficients are zero (see Appendix A for further details). The full models with covariates describe the data better than simple models without covariates.

6.1.1 Survivor Curves for Select Transitions

It is often easier to interpret the estimation results by plotting survivor functions. Survivor curves give the probability that the household remains in the beginning state at some time t given that the household has not made a transition before t. So, for the transition from single to couple, the survivor curve gives the probability that the person has not yet married.

Survivor curves depends on the covariate values, some of which may be time-varying. To aid in the presentation, I have selected a few representative households and transitions. The following figures show survivor curves for the transitions from 'Single' to 'Single with 1 child', from 'Single' to 'Couple', from 'Couple with 2 children' to 'Single with 2 children', and from 'Couple' to 'Couple with 1 child'. Only a few representative households were chosen, but others may be added later.

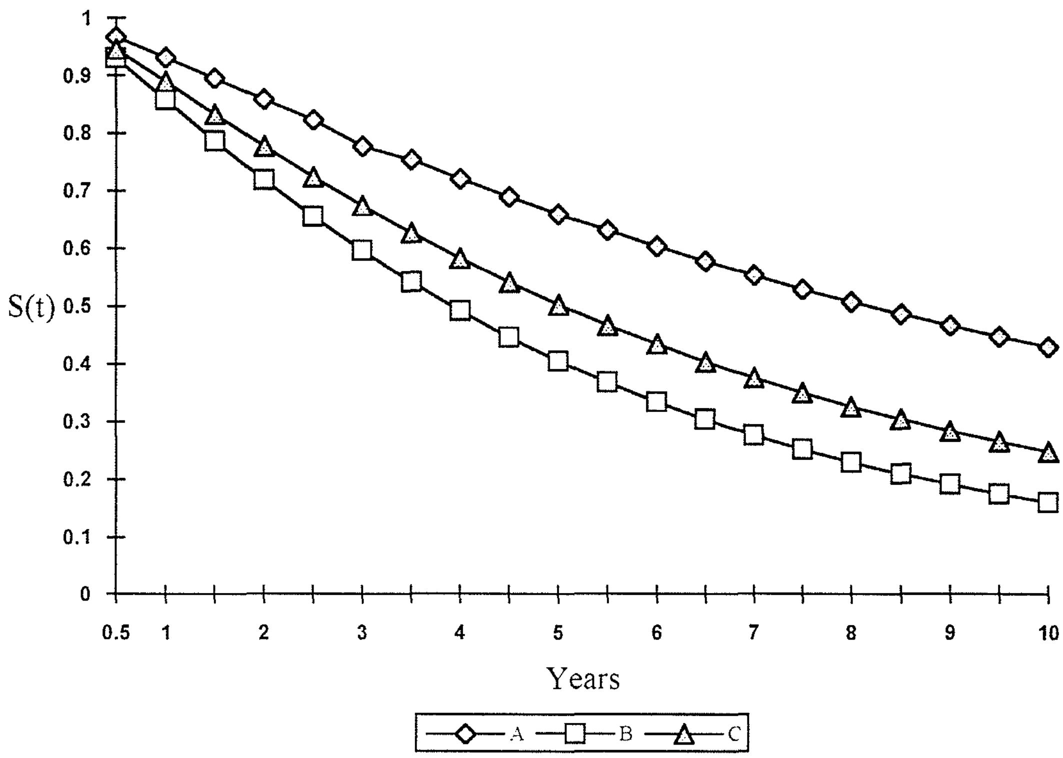
Fig 2. Survival Curves for Single to Single with 1 child



A = Black female who dropped out of high school, is employed, earns $15,000/year ($7500 per 6 months), and was age 18 in year 1.
B = White female who dropped out of high school, is employed, earns $15,000/year, and was age 18 in year 1.
C = White female who is a college graduate, is employed, earns $30,000/year, and was age 22 in year 1.

As this figure shows, black women who drop out of high school are most likely to have a child without being married. Their survival curve slopes down the quickest, implying that they transition from single to single with a child sooner. For both cateogies of white women, the chance of survival is much higher, e.g. they are less likely to have a child before marriage. But differences still exist between highly educated, higher earning white women and their less educated, poorer counterparts. Those who dropped out of high school and earn less are likely to transition sooner.

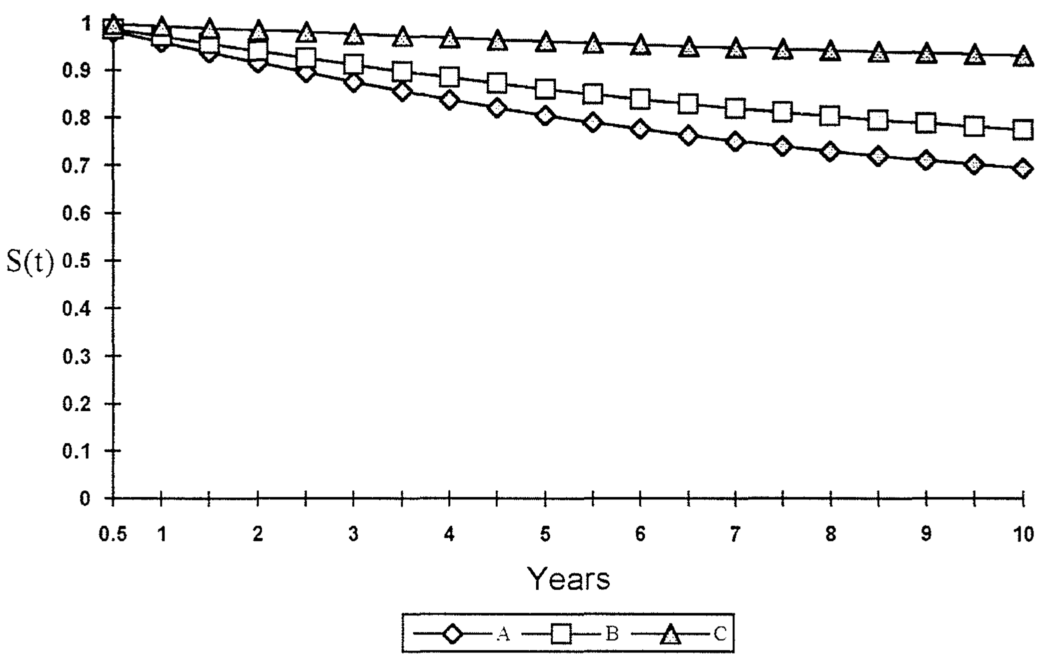
Fig. 3 Survival Curves for Single to Couple



A = Black female who dropped out of high school, is employed, earns $15,000/year ($7500 per 6 months), was age 18 in year 1, and was not previously married.
B = White female who dropped out of high school, is employed, earns $15,000/year, was age 18 in year 1, and was not previously married.
C = White female who is a college graduate, is employed, earns $30,000/year, was age 22 in year 1, and was not previously married.

The survival curves for these three groups of women are much more similar than the previous figure. White women who drop out of high school are the most likely to marry at an earlier age. The next most likely to marry early are black women with little education. Finally, white women with college degrees are the most likely to postpone marriage. Keep in mind that I have included long-term cohabitation relationship in the married category.

Fig. 4 Survival Curves for Couple with 2 children to Single with 2 children



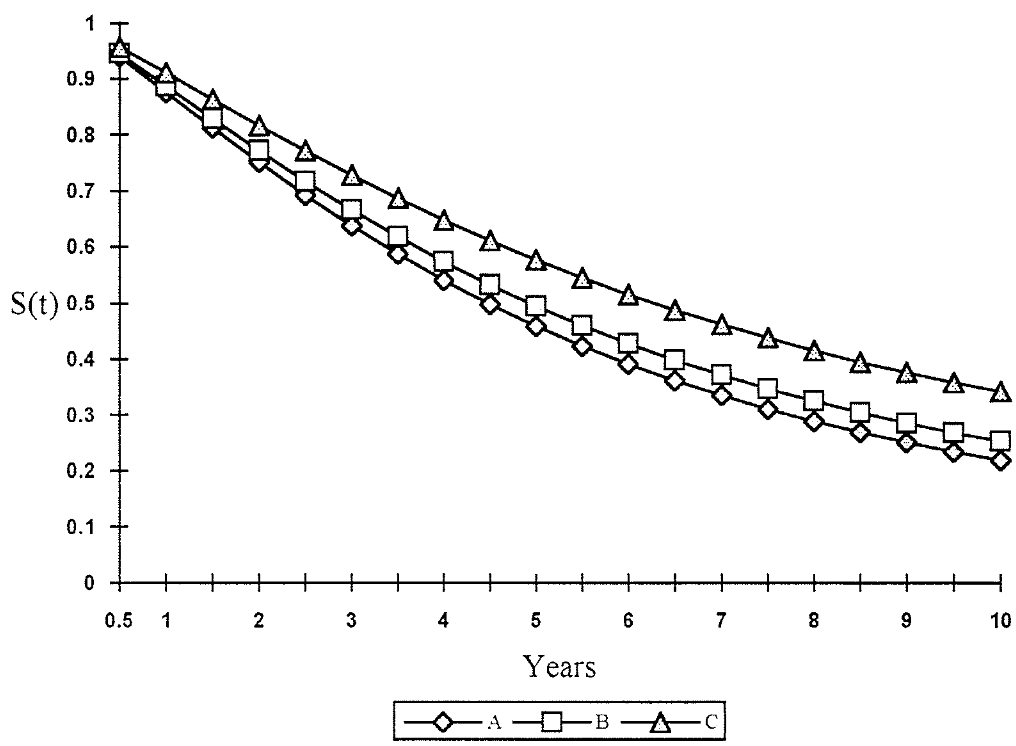
A=Black household, husband and wife both employed, household income of $15,000/year, both husband and wife dropped out of high school and were age 25 in year 1. Both children between the ages of 6 and 18. Children were not born before the marriage.

B=White household, husband and wife both employed, household income of $15,000/year, both husband and wife dropped out of high school and were age 25 in year 1. Both children between the ages of 6 and 18. Children were not born before the marriage.

C=White household, husband employed, wife is out of the work force, household income of $40,000/year, both husband and wife went to college and were age 30 in year 1. Both children between the ages of 6 and 18. Children were not born before the marriage.

This figure shows that blacks are likely to transition from 'Couple with 2 children' to 'Single with 2 children' before whites. The household most likely to remain in the stable marriage consists of a housewife (e.g. she is out of the workforce) and a sole breadwinning husband who makes a relatively high income ($40,000/year).

Fig. 5 Survival Curves for Couple to Couple, 1 child



A=Black household, husband and wife are both employed, household income of $15,000/year, both husband and wife dropped out of high school and were age 25 in year 1.
B=White household, husband and wife are both employed, household income of $15,000/year, both husband and wife dropped out of high school and were age 25 in year 1.
C=White household, husband and wife are both employed, household income of $40,000/year, both husband and wife went to college and were age 30 in year 1.

This figure shows that the timing of child birth is similar for these three categories of households. Differences across racial groups are not as dramatics as some of the previous figures. Wealthier households are more likely to postpone child birth. What this figure does not show is that households which consists of a stay-at-home wife are likely to have a child before households with working women.

6.2 Unobserved Heterogeneity

Next I have grouped transitions according to possible inter-relations in the hazard functions. As mentioned earlier, the decision to have another (or a first) child may be predicated by the chance of divorce, and vice verse. These types of inter-relationships may be captured with unobserved heterogeneity which will be estimated within the following groups shown in Table 13.

Table 13: Groups of Transitions for Unobserved Heterogeneity Estimation

Group 1:	Group 2:
Single → Couple	Couple, 1 child → Single
Single → Couple, 1 child	Couple, 1 child → Couple
Single → Single, 1 child	Couple, 1 child → Couple, 2 children
Single → Other	Couple, 1 child → Single, 1 child
Couple → Single	Couple, 2 children → Single
Couple → Couple, 1 child	Couple, 2 children → Couple, 1 child
Other → Single	Couple, 2 children → Single, 2 children
Other → Couple	
Other → Couple, 1 child	
Group 3:	Group 4:
Single, 1 child → Single	Couple, 2 children → Couple, 3+ children
Single, 1 child → Couple, 1 child	Couple, 3+ children → Couple, 2 children
Single, 1 child → Single, 2 children	Couple, 3+ children → Single, 3+ children
Single, 2 children → Couple, 2 children	Single, 2 children → Single, 3+ children
Single, 2 children → Single, 1 child	Single, 3+ children → Single, 2 children

Estimation results for all groups indicate that unobserved heterogeneity is insignificant. When important explanatory variables were excluded (for example, race), unobserved heterogeneity was estimated to be significant. Therefore, the full models with all covariates are complete in the sense that no significant unobserved heterogeneity remains.

7 Further Research

The models must be used to actually forecast demographic transitions. Further research will focus on the implementation of that forecasting procedure. Without unobserved heterogeneity, forecasting is rather straightforward and uses the survival curves for all possible transitions out of the current states (e.g. singles can transition to single with 1 child, couple, couple with 1 child, and other). At each time period, there is some probability with each transition. A random variable can be drawn to indicate which transition will actually occur. New household characteristics must be updated to account for the transition.

The procedure would not be so straightforward if unobserved heterogeneity had been significant. Either I would need to add additional variables to remove the unobserved heterogeneity, or simulate with it included. To simulate with unobserved heterogeneity, I would need to draw a random component to be added to each household's calculated hazard functions. The random component would come from a normal distribution with mean 0 and variance determined by the estimated parameter C_{ij} from equation 12.

Finally, I would like to estimate these models using individual level data with the same household descriptions as states. Duncan and Hill (1985) suggest this modeling strategy, and I think it would be useful to compare final microsimulation results using both the current strategy and the individual-level strategy.

APPENDIX A: LIKELIHOOD RATIO TEST STATISTICS

The log of the likelihood ratio gives the following test statistic:

$$-2\left[\text{Log likelihood } (\hat{\gamma}_0, \hat{\gamma}_1) \text{ - Log likelihood } (\hat{\gamma}_0, \hat{\gamma}_1, \hat{\beta})\right]$$

It follows a χ^2 distribution with degrees of freedom equal to the dimension of $\hat{\beta}$ (or the number of covariates included in the full model which are restricted to zero in the basic model without covariates). Table B.1 shows the value of the log likelihood test statistic and the corresponding cutoff value from the χ^2 distribution with appropriate degrees of freedom.

Table B1: Likelihood Ratio Tests Statistics

	$-2[LL(\hat{\gamma}_0, \hat{\gamma}_1) - LL(\hat{\gamma}_0, \hat{\gamma}_1, \hat{\beta})]$	$\chi_n^2(0.995)$
Single → Couple	408.02	28.30
Single → Couple. 1 child	145.22	28.30
Single → Single. 1 child	120.20	26.76
Single → Other	108.36	26.76
Single. 1 child → Single	119.06	28.30
Single. 1 child → Couple. 1 child	156.22	29.82
Single. 1 child → Single. 2 children	24.34*	28.30*
Single. 2 children → Couple. 2 children	59.86	29.82
Single. 2 children → Single. 1 child	79.62	29.82
Single. 2 children → Single. 3+ children	63.78	29.82
Single. 3+ children → Single. 2 children	48.18	29.82
Other → Single	142.42	26.76
Other → Couple	78.66	26.76
Other → Couple. 1 child	35.92	25.19
Couple → Single	467.00	35.72
Couple → Couple. 1 child	1067.22	35.72
Couple. 1 child → Single	128.34	38.58
Couple. 1 child → Couple	627.02	37.16
Couple. 1 child → Couple. 2 children	589.30	37.16
Couple. 1 child → Single. 1 child	193.68	38.58
Couple. 2 children → Single	70.38	40.00
Couple. 2 children → Couple. 1 child	644.94	38.58
Couple. 2 children → Couple. 3+ children	185.46	38.58
Couple. 2 children → Single. 2 children	167.90	40.00
Couple. 3+ children → Couple. 2 children	351.14	40.00
Couple. 3+ children → Single. 3+ children	129.54	41.40

* Does not pass the likelihood ratio test. Simple model without covariates is chosen over the more complex model.

References

Becker, Gary S., landis, Michael and Robert Michael (1977): "An economic analysis of marital instability", *Journal of Political Economy*, Vol. 85, pp. 1141-1187.

Bennet, Neil G., Bloom, David E. and Patricia H. Craig (1989): "The Divergence of Black and White Marriage Patterns", *American Journal of Sociology*, Vol. 95 No. 3, pp. 692-722.

Blau, Francine D. and Marianne A. Ferber (1986): *The Economics of Women, Men and Work*, Prentice-Hall, Englewood Cliffs, New Jersey.

Bloom, David E. and Anne R. Pebley (1982): "Voluntary Childlessness: A Review of the Evidence and Implications", *Population Research and Policy Review 1*, pp. 203-224.

Bloom, David E. and James Trussell (1984): "What are the Determinants of Delayed Childbearing and Permanent Childlessness in the United States?", *Demography*, Vol. 21 No. 4, pp. 591-611.

Bumpass, Larry L. (1990): "What's Happening the the Family? Interactions Between Demographic and Institutional Change", *Demography*, Vol. 27 No. 4, pp. 483-498.

Bumpass, Larry L., Sweet, James and Andrew Cherlin (1991): "The Role of Cohabitation in Declining Rates of Marriage", *Journal of Marriage and the Family*, Vol. 53 No. 4, pp. 913-927.

Bumpass, Larry L., Sweet, James and Teresa Castro Martin (1990): "Changing Patterns of Remarriage", *Journal of Marriage and the Family*, Vol. 52 No. 3, pp. 747-756.

California Air Resources Board (CARB) (1994): "Zero-Emission Vehicle Update: Draft Technical Document for the Low-Emission Vehicle and Zero-Emission Vehicle Workshop", CARB, Sacramento, March.

Chang, H. L. and P. Jovanis (1990): "Formulation accident occurrence as a survival process", *Accident Analysis and Prevention*, Vol. 22 No. 5, pp. 407-419.

Cherlin, Andrew J. (1992): *Marriage, Divorce, Remarriage*, Cambridge, MA, Harvard University Press.

Cohanim, S., M. Hoggan, R. Sin, L. W. Jones, M. Hsu, S. Tom (1991): *Summary of Air Quality in California's South Coast and Southeast Desert Air Basin, 1987-1990*, South Coast Air Quality Management District, August 1991.

Duncan, G. J. and M. S. Hill (1985): "Conceptions of Longitudinal Households", *Journal of Economic and Social Measurement*, Vol. 13, pp. 361-375.

Ermisch, J. (1988): "An Economic Perspective on Household Modelling", Chapter 3 in *Modelling Household Formation and Dissolution*, edited by Nico Keilman, Anton Kuijsten and Ad Vossen, Clarendon Press, Oxford, pp. 23-39.

Flinn, C. J. and J. J. Heckman (1982): "Models for the Analysis of Labor Force Dynamics", *Advances in Econometrics*, Vol. 1, pp. 35-95.

Flinn, C. J. and J. J. Heckman (1983): "Erratum and Addendum to Volume 1: Models for the Analysis of Labor Force Dynamics ", *Advances in Econometrics*, Vol. 2, pp. 219-223.

Glick, Paul C. (1990): "American Families: As They Are and Were", *Sociology and Social Research*, Vol. 74 No. 3, pp. 139-145.

Glick, Paul C. and Sung-Ling Lin (1987): "Remarriage After Divorce, Recent Changes and Demographic Variations", *Sociological Perspectives*, Vol. 30 No. 2, pp. 162-179.

Glick, Paul C. and Arthur J. Norton (1977): "Marrying, Divorcing, and Living Together in the U. S. Today", *Population Bulletin*, pp. 3-39.

Hall, Jane, Arthur Winer, Michael Kleinman, Frederick Lurmann, Victor Brajer, and Steven Colome (1992): "Valuing the Health Benefits of Clean Air", *Science*, Vol. 225, pp. 812-817.

Heaton, Tim B. and Cardell K. Jacobson (1994): "Race Differences in Changing Family Demographics in the 1980s", *Journal of Family Issues*, Vol. 15, No. 2, pp. 290-308.

Hensher, David A. (1994): "The Timing of Change for Automobile Transactions: A Competing Risk Multispell Specification", working paper presented at the International Conference on Travel Behavior, Santiago, Chile, June 1994.

Hensher, David A. and Fred L. Mannering (1994): "Hazard-based Duration Models and Their Application to Transport Analysis", *Transport Reviews*, Vol. 14 No. 1, pp. 63-82.

Hensher, D., N. C. Smith, F. W. Milthorpe, and P. O. Barnard (1992): *Dimensions of Automobile Demand*, Elsevier, Amsterdam.

Kalbfleisch, J. D. and R. L. Prentice (1980): *The Statistical Analysis of Failure Time Data*, Wiley Publishing, New York.

Krupnick, Alan and Paul Portney (1991): "Controlling Urban Air Pollution: A Benefit Cost Assessment", *Science*, Vol. 252, pp. 522-528.

Lancaster, T. (1990): *The Econometric Analysis of Transition Data*, Cambridge University Press.

Lancaster, Tony (1979): "Econometric Methods for the Duration of Unemployment", *Econometrica*, Vol. 47 No. 4, pp. 939-956.

Land, Kenneth and Andrei Rogers (1982): *Multidimensional Mathematical Demography*, proceedings from the Conference on Multidimensional Mathematical Demography, University of Maryland, March 1981, edited by Kenneth Land and Andrei Rogers, Academic Press, Inc.

Lillard, Lee A. (1993): "Simultaneous Equations for Hazards", *Journal of Econometrics*, Vol. 56, pp. 189-217.

Martin, Teresa Castro and Larry L. Bumpass (1989): "Recent Trends in Marital Disruption", *Demography*, Vol. 26 No. 1, pp. 37-51.

Merz, Joachim (1991): "Microsimulation - A Survey of Principles, Developments and Applications", *International Journal of Forecasting*, Vol 7, pp. 77-104.

Meyer, Bruce D. (1990): "Unemployment Insurance and Unemployment Spells", *Econometrica*, Vol. 58 No. 4, pp. 757-782.

Moore, Kristin A. and Thomas M. Stief (1991): "Changes in Marriage and Fertility Behavior, Behavior Versus Attitudes of Young Adults", *Youth & Society*, Vol. 22 No. 3, pp. 362-386.

Morgan, S. Phillip, Lye, Diane N. and Gretchen A. Condran (1988): "Sons, Daughters, and Risk of Marital Disruption", *American Journal of Sociology*, Vol. 94 No. 1, pp. 110-129.

Parrish, M. (1994): "EPA to Toughen Emissions Rules in the Northeast", *Los Angeles Times*, The Times Mirror Company, September 14, 1994, pp. D1.

Pyke, Karen D. (1994): "Women's Employment as a Gift or Burden, Marital Power Across Marriage, Divorce, and Remarriage", *Gender & Society*, Vol. 8 No. 1, pp. 73-91.

Ren, Weiping, David Brownstone, David Bunch, and Thomas Golob (1994): "A Vehicle Transactions Choice Model For Use in Forecasting Demand for Alternative-Fuel Vehicles", submittted to *Research in Transportation Economics*.

Rima and Van Wissen (1987): *A Dynamic Model of Household Relocation, A Case Study for the Amsterdam Region*, Free University Press, Amsterdam.

Rinduss, Ronald R. and Craig St. John (1983): "Social Determinants of Age at First Birth", *Journal of Marriage and the Family*, pp.553-565.

South Coast Air Quality Management District (SCAQMD) and Southern California Association of Governments (SCAG) (1994): *Air Quality Management Plan: Meeting the Clean Air Challenge*, SCAQMD, Diamond Bar, CA, August.

Spanier, Graham B. and Paul C. Glick (1981): "Marital Instability in the United States: Some Correlates and Recent Changes", *Family Relations*, Vol. 31, pp. 329-338.

Stanley, Sandra C., Hunt, Janet G. and Larry L. Hunt (1986): "The Relative Deprivation of husbands in Duel-Earner Households", *Journal of Family Issues*, Vol. 7 No. 1, pp. 3-20.

Thorton, Arland (1988): "Cohabitation and Marriage in the 1980s", *Demography*, Vol. 25 No. 4, pp. 497-508.

Treachman, Jay D. and Mark D. Hayward (1993): "Interpreting Hazard Rate Models", *Sociological Methods & Research*, Vol. 21 No. 3, pp. 340-371.