Trends in Childlessness Among Less Educated and Minority Women in the United States.

Steven P. Martin

Levels of childlessness among U.S. women age 40 to 44 have almost doubled in the last two decades, to about 19 percent. During this time, social concern over childlessness and academic investigations of childlessness have focused on the highly successful, career-oriented women who have recently been the most likely to remain childless. However, many of the social concerns related to childlessness are at least as salient among economically disadvantaged groups of women as among their more economically successful counterparts. For one example, a lack of family support at old age can be a particular problem for women without substantial assets and economic resources. For another example, diminishing future fertility prospects force many women to make hard choices about jobs and family, but those choices may be even harder for socioeconomically disadvantaged groups of women whose health declines more rapidly with age and whose marriage market options in early adulthood may be bleak.

It is difficult to project fertility patterns with distinctions by racial and educational groups, and some notable demographic projections have turned out to be inaccurate. This problem may be compounded for recent birth cohorts of women because the typical family paths that lead to childlessness among less educated and minority women may be distinct from the pathways to childlessness among highly educated, economically successfully women. Family trends among highly educated and white women have been distinguished by postponement of marriage and fertility, with declining marital fertility in young adulthood partially offset increases in marriage and fertility at later ages. In comparison, family trends among less educated and particularly black women have been distinguished by delays or declines in marriage, with declining rates of marital first births partially offset by increases in nonmarital fertility.

This paper is a methodological exercise that borrows an existing, straightforward fertility projection technique used by Rindfuss, Swicegood, and Morgan (1988) and improves it by incorporating competing risks for first birth and marriage transitions. In other words, I will estimate competing hazard models for first marriage and nonmarital first births, as well as models for marital first births and disruptions of childless marriages, then use age-specific hazard rates to reconstruct the proportion of women remaining childless to age 45, for different education and racial groups. The data source will be the marriage and fertility histories in the 2001 Survey of Income and Program Participation (SIPP), supported by data from marriage and fertility histories in June Current Population Surveys for 1990 and 1995.

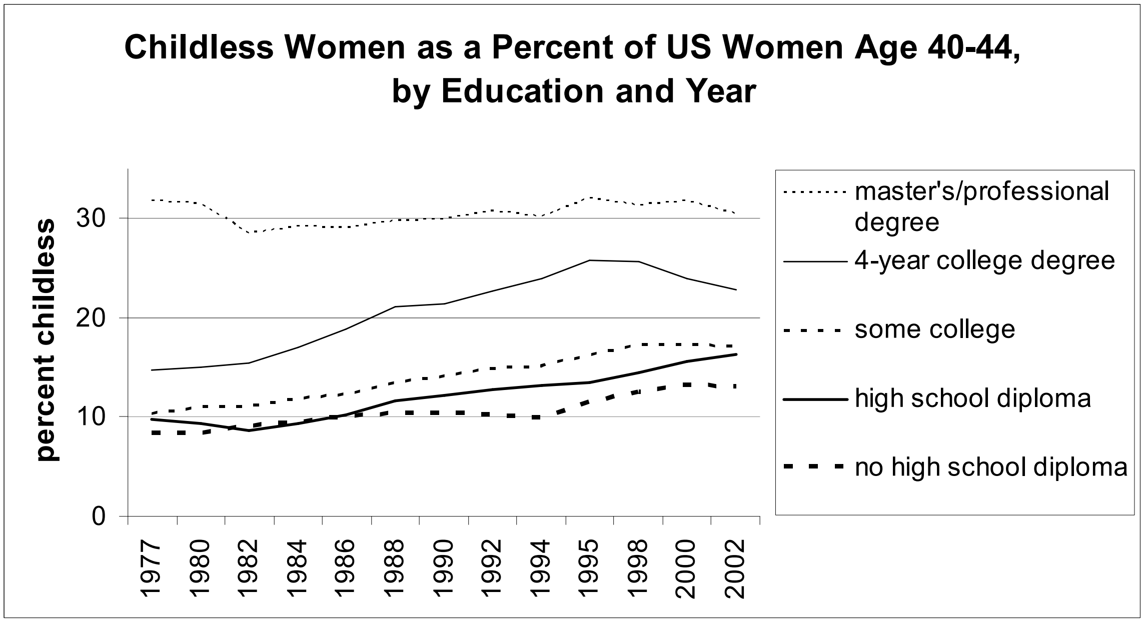
Here are some preliminary descriptive data from the fertility supplements to the 1977 – 2002 June Current Population Surveys (CPS). Figure 1 shows that rates of childlessness remain lower among less educated women than among their more educated counterparts, but the gap appears to be closing, particularly since the mid-1990s. One objective of this paper is to determine whether the gap in childlessness might continue to close in the near future.

< Figure 1 about here >

Figure 2 shows preliminary descriptive data, also from the June CPS series, but instead of showing childlessness as a percentage of each educational group, it shows the estimated number of women childless, as well as the marital status of those women. Figure 2 demonstrates that most childless women have less than a college degree, just as most women have less than a college degree, and that a large proportion of women without a college degree who are childless are also not married. Hence, an analysis that explicitly models first births in the context of transitions into and out of marriage is not only tapping an underutilized source of information, but a characteristic of childlessness that may have important policy implications in its own right.

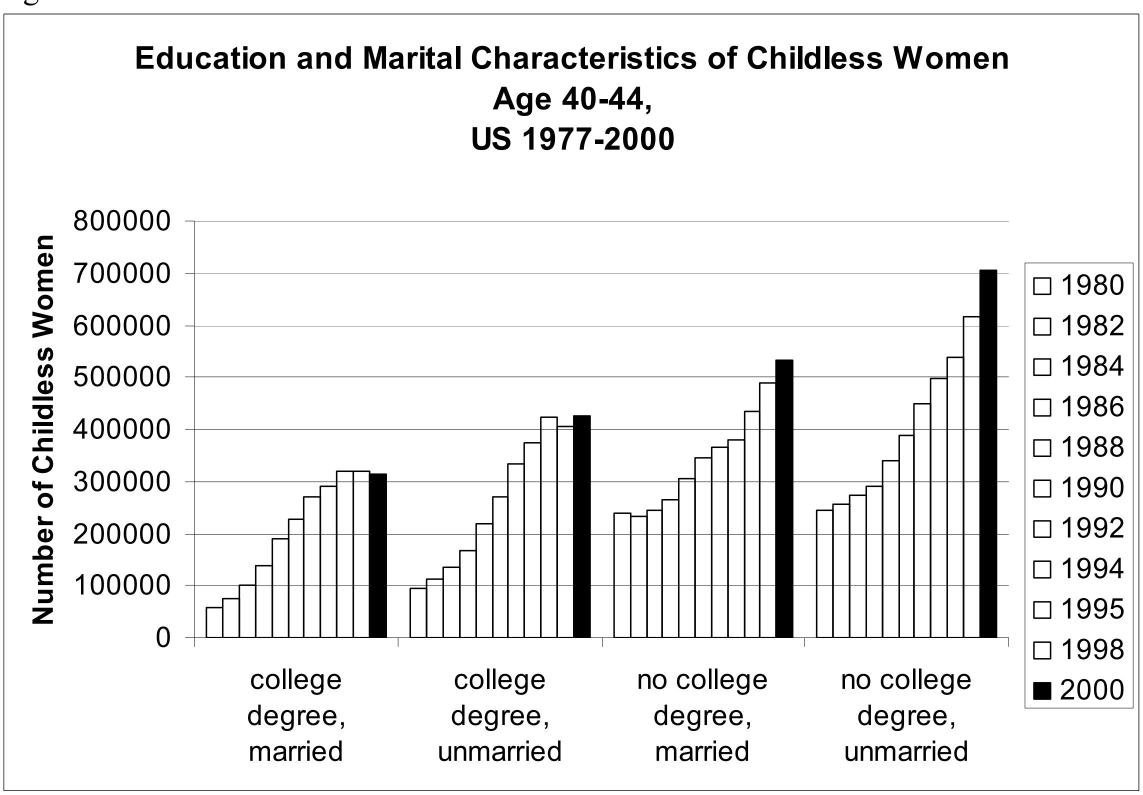
< Figure 2 about here >

Figure 1.



Source: June Current Population Survey series.

Figure 2.



Source: June Current Population Survey series.

Incomplete Technical Appendix.

(Here follows a description of the procedure for projecting first births in the absence of competing risks of a first marriage. I am still writing a formal description of the procedure when competing risks are used.)

Table 2 portrays the age-specific survivor probabilities for first birth for a series of birth cohorts. In this table, $S_{c,t}$ represents the proportion of individuals in a five-year birth cohort *C* who remain childless to the end of age interval *T*, adjusted for any censoring during age interval *T* using a Kaplan-Meyer life table procedure. The objective of the estimation procedure is to project the missing cell values for birth cohorts 3, 4, and 5 (with increasing uncertainty as the number of missing cells increases for more recent cohorts.

< Table 2 about here >

I begin by demonstrating the estimation procedure for a cohort with no missing data. Cohorts 1 and 2 in Table 2 already have more or less complete information on birth from age 15 to 44. For a cohort *C* with complete data across all age groups to age 45, the proportion P_C childless at age 45 is calculated by equation 1. In equation 1, - $\Delta S_{C,t}$ is the proportion of persons in cohort *C* experiencing a first birth in interval *t*, adjusted for any censoring within the cohort using Kaplan-Meier life table estimation.

$$P_C = \prod_{t=15}^{44} \left(\frac{S_{C,t} - \Delta S_{C,t}}{S_{C,t}} \right) = S_{C,45} \qquad (1)$$

Next I demonstrate a procedure for cohorts with missing data. Cohorts 3 through 5 in Table 3 do not have complete data to age 45. The proportion childless cannot be estimated directly for these cohorts, so one uses the latest available data from cells on the left to fill missing cells. Given a 5-year birth cohort *C* with valid observations of age-specific survivorship up to age T_1, the previous 5-year cohort *C-1* with valid observations of age-specific survivorship up to age T_2, and all preceding 5-year cohorts *C-n* with valid observations of age-specific survivorship up to age T_n, up to the point where $T_n >= 44$, the projected proportion of individuals in cohort *C* childless to age 45 is as follows:

$$S_{C,45} = \prod_{t=15}^{T_1}\left(\frac{S_{C,t} - \Delta S_{C,t}}{S_{C,t}}\right) \prod_{t=T_1+1}^{T_2}\left(\frac{S_{C-1,t} - \Delta S_{C-1,t}}{S_{C-1,t}}\right) \ldots \prod_{t=T_{n-1}+1}^{44}\left(\frac{S_{C-n,t} - \Delta S_{C-n,t}}{S_{C-n,t}}\right) \qquad (2)$$

One can modify this procedure to allow for unmeasured heterogeneity in the childbearing population. For each successive cohort, assume some unknown proportion of the population $S_{(s)}$ is unable or unwilling to marry at any age. The predicted proportion of individuals in cohort *C* not marrying by age 45 is then:

$$S_{(s)} + S_{C,45} = S_{(s)} + \prod_{t=15}^{T_1}\left(\frac{S_{C,t} - S_{(s)} - \Delta S_{C,t}}{S_{C,t} - S_{(s)}}\right) \prod_{t=T_1+1}^{T_2}\left(\frac{S_{C-1,t} - S_{(s)} - \Delta S_{C-1,t}}{S_{C-1,t} - S_{(s)}}\right)$$

$$\ldots \prod_{t=T_{n-1}+1}^{44}\left(\frac{S_{C-n,t} - S_{(s)} - \Delta S_{C-n,t}}{S_{C-n,t} - S_{(s)}}\right) \qquad (3)$$

When one uses a spreadsheet program to solve equations such as this, one encounters two difficulties. Firstly, one can only adjust for unmeasured heterogeneity of the finite mixture or “mover-stayer” form. To evaluate whether alternate forms of unmeasured heterogeneity affect the estimates, I used the aML statistical program (Lillard and Panis 2000) to replicate the main models with an assumed normally distributed unmeasured heterogeneity factor. The results were substantively the same. Secondly, while unmeasured heterogeneity is present in every population, it is difficult to identify. Social scientists have developed various techniques to statistically control for it (see Heckman and Singer 1984; Palloni and Sorensen 1990; Blossfeld and Hamerle 1992), but it is usually impossible to identify the form and level of unmeasured heterogeneity without explicit and often unsupportable assumptions. Instead of trying to identify the level of unmeasured heterogeneity in a given population, I solve equation 3 for a range of possible values of $S_{(s)}$.

As a next, step, I adjust the procedure to allow for changes in age-specific birth rates. Instead of using observed age-specific cohort survivorship from previous cohorts, I identify the linear trend in age*cohort survivorship using ordinary least squares regression. The predicted proportion of individuals in cohort *C* not marrying by age 45 is described by equations 4 through 6.

$$S_{C,45} = \prod_{t=15}^{T_1} \left(\frac{S_{C,t} - \Delta S_{C,t}}{S_{C,t}} \right) * f(T_1 + 1, T_2) \ldots * f(T_{(N-1)} + 1, T_N) \quad (4)$$

where the proportion remaining childless through an age interval is a function of the cohort *C*

$$f(T_{(n-1)} + 1, T_n) = b_0 + b_1 C \quad (5)$$

and where b_0 and b_1 are estimated by ordinary least squares regression using the equation

$$\prod_{t=T_{n-1}}^{T_n}\left(\frac{S_{c,t} - \Delta S_{c,t}}{S_{c,t}}\right) = b_0 + b_1 c + e \qquad (6)$$

for values of c from c = 1 to c = C-1

Finally, I combine the adjustments for linear trends from equations 4 through 6 with the controls for mover-stayer heterogeneity from equation 3. The resulting algorithm has the following form:

$$S_{C,45} = \prod_{t=15}^{T_1}\left(\frac{S_{C,t} - \Delta S_{C,t} - S_s}{S_{C,t} - S_s}\right) * f(T_1 + 1, T_2) \ldots * f(T_{(N-1)} + 1, T_N) \qquad (7)$$

with the following adjustment to the ordinary least squares regression equation for b_0 and b_1:

$$\prod_{t=T_{n-1}}^{T_n}\left(\frac{S_{c,t} - \Delta S_{c,t} - S_s}{S_{c,t} - S_s}\right) = b_0 + b_1 c + e \qquad (8)$$

I provide confidence intervals for these estimation procedures in two ways. First, I estimate models across a range of assumptions about the levels of heterogeneity in the population and the presence or absence of linear trends in age-specific birth rates. Then, for each model based on its own set of assumptions, I use bootstrap procedures to draw 120 random samples with replacement and assess the 95% confidence interval for the 120 samples. Full details of the bootstrap procedure are available on request.