DataFirst Technical Papers

DataFirst

Grouped-Data Identification of Marital Behaviour from Repeated Cross-sections in South Africa

by
Grace Kumchulesi

Technical Paper Series
Number 6

Recommended citation

Kumchulesi, G. (2008). Grouped-Data Identification of Marital Behaviour from Repeated Cross-sections in South Africa. A DataFirst Technical Paper Number 6. Cape Town: DataFirst, University of Cape Town

DataFirst, University of Cape Town, Private Bag, Rondebosch, 7701, Tel: (021) 650 5708, Email: info@data1st.org/support@data1st.org

Grouped-Data Identification of Marital Behaviour from Repeated Cross-sections in South Africa

Grace Kumchulesi[1]

Abstract

South Africa, like most developing countries, suffers from lack of complete vital statistics recording system. Nevertheless, regular collection of nationally represented data, with the inception of annual October Household Surveys in 1993 help to indirectly measure the country's dynamic processes. New samples are drawn from these independent surveys so that it is possible to follow cohorts of individuals over time. It is of interest, therefore, to examine whether the various changes observed are real or result from differences in the sampling designs. This study uses the sample age-cohort means to identify marital behaviour of prime-aged indigenous women in South Africa. Simple graphical techniques are used for this purpose. The study identifies a shift in marriage patterns between the younger and older cohorts, with a higher proportion of the older cohorts marrying than the younger cohorts. Further evidence from the decomposition analysis using age-period-cohort models indicate that period effect is constant around zero across the period under study while the age and cohort effects are non-zero and significant, suggesting that the observed shift in marital behaviour is real.

Key words: APC models, Synthetic/Pseudo Panel, Marriages

1.0 Introduction and Motivation

South Africa, like most developing countries, suffers from lack of panel data records of vital statistics, which makes studying important dynamic processes almost impossible. Nevertheless, the availability of nationally represented datasets collected across several cross-sections enables researchers to undertake dynamic studies by tracking groups of individuals, known as cohorts, through such data (Deaton, 1985). Due to its transitional nature in an individual's life history, which is critical, and the central role it has on family, how marital behaviour evolves overtime is one issue that has attracted research among social scientists, including demographers, sociologists, anthropologists and economists, yet not much research has been done on this issue in Africa.

In Africa, marriage is typically universal, characterised by early entry and high proportions of individuals married by the age of 40 (Adegboga, 1994). In South Africa, however, several studies (including Kalule-Sabiti et al., 2007; Posel, 2003; Casale and Posel, 2002) have noted a change in marriage pattern for both men and women. The change is most prominent among women, and especially African[2] women. The phenomenal nature of the changes in marital behaviour is evident in the post-apartheid South Africa[3] and portrays a pattern of marriage postponement, which in some cases is indefinite. For registered marriages in South Africa, for example, Statistics South Africa (2002a) reports that the median age at first marriages in 1999 was 33.7 years for males and 29.5 years for females[4]. Casale and Posel (2002) note a decline in married proportions from 38.7 percent in 1995 to 34.5 percent in 1999 for African women of working age (between 15 and 59) but the decline was largest in the ages between 25 and 34 years. Ntuli (2007) notes a further decline to 31.4 percent in 2004. Thus, trends in the marriage pattern in South Africa point to both delayed marriages and declining marriage rates.

However, even though comparing proportions of individuals ever married between time points (or overtime) may give an indication on the question of whether there are fewer

---

[2] According to the racial categorization in South Africa, blacks include Indians, Coloureds and indigenous South Africans. Indigenous South Africans are called Africans, to separate them from the rest of the black population, and, according to Census 2001, they make about 80 percent of South Africa's population.
[3] South Africans went to the polls in the spring of 1994 and realized a new political regime which was democratically elected.
[4] According to a proposition of the timing at first marriage for females by Bogue (1969), late marriage is defined by median age at first marriage of 22 years and over. For further evidence of late marriage pattern in South Africa, see Kalule-Sabiti, (2007).

marriages, and may be equally instructive to calculate age at first marriage to determine if the marriage pattern portrays later marriages, the observed change may not be a real change in marital behaviour. Especially when dealing with independent survey data, care has to be taken as far as investigation of generational changes is concerned. Factors related to sampling design, which may be different for different surveys, may contribute to the observed pattern. For example, in their study of dynamics in the labour market in South Africa between 1994 and 2004, Branson and Wittenberg (2007) noted that the higher levels of employment in 2000, relative to the preceding years may be due to the increased attention by Statistics South Africa, to find informal activities. This yields the concern that observed decline in marriages may be a reflection of sampling issues rather than a real change in marital behaviour. It is against this concern, that this study is motivated to establish whether the recent changes documented for marriage pattern is real or not.

Dynamic analysis is essential in understanding behaviour. With the absence of authentic panel data capturing marriage transitions in South Africa, this study makes use of synthetic panel data constructed from a series of independent cross-sections. We will track cohorts through these datasets to identify marital behaviour among African women. A cohort in this study is defined by year of birth, which is equivalent to age in 1995. Simple graphic analyses will be employed to describe the recent trends in marriage for indigenous South African women. Further, a decomposition analysis will be undertaken to establish the source of the change in marital patterns.[5]

Use of synthetic panel data in place of an authentic panel data does not translate into obtaining inferior results. If anything, the synthetic panel does not suffer from most of the weaknesses faced by a genuine panel. For instance, attrition bias, which may result from refusal, death or morbidity in a real panel, is absent in a synthetic panel. Representativeness is constantly maintained in synthetic panels since new samples are drawn every year (Deaton, 1997).

---

[5] Equally striking change paralleling decline in marriages is the increase in labour force participation among indigenous African women. Interdependency issues between marriage and labour supply decisions will therefore, as may be anticipated, be the next issue of concern. In other words, we will seek to establish the relationship between the two, with their joint determination characteristic taken into account. Further, it will be of general interest to investigate Gary Becker's efficient marriage market hypothesis where issues of assortative mating and marginal product pricing will be addressed. In this regard, this paper, therefore, serves as a motivational paper and point of departure, providing a big picture of marriage trends in South Africa before the micro-picture is investigated in the two papers that follow.

Use of cohort data has its weaknesses too. Two main disadvantages are loss of individual information due to aggregation of information and the assumption of constant population. The later might not be realistic since distributions of variables in a population tend to change over time resulting from aging, migration and death (Deaton, 1997). As a result, the identical distribution assumption is not always valid, but the independence assumption is. The observations in the pooled cross sections over time are independent but not identically distributed (i.n.i.d). However, these weaknesses do not necessarily invalidate use of cohort data, especially where individual-level panel data is not available. Where cohorts have been defined to ensure homogeneity within cohorts, the former problem may be avoided. The later problem may be avoided by restricting the cohorts to some age range. We take care of the later problem by sampling prime-aged women in the age group 15 to 49.

The rest of section 1 contains the problem statement and significance of the study, as well as the research objectives. Section 2 reviews the theoretical and empirical literature of marriage. Section 3 describes the data while section 4 contains methodological issues. In section 5, we will present the results which show that the change in marital behaviour is a real change. The remainder of section 5 explores descriptively some of the possible reasons for this shift in marriage behaviour over the period. In particular, marriage is examined for different cohorts in a univariate context by education level and labour force participation. Marriage market factors in the form of sex ratios are also examined. While the focus will be on comparing cohorts of younger and older generations of African women, data on African men as well as white women and men will also be provided. This comparison will not only highlight differences in marital behaviour between the two African women generations, but will also provide an indication of the position of average marriages across gender and racial lines over the period. Section 6 concludes.

1.1 Problem Statement and Significance of the Study

As depicted in the preceding section, the problem identified to motivate the study is the observed decline and delay in marriage in the post-apartheid South Africa. The question that this study attempts to answer is whether the observed change in marriage pattern is a generational change in marital behavior. Of fundamental importance in the analysis is to determine whether the change in marriage patterns is cohort-based or based on some other factors such as age or calendar year. This makes it worthwhile to identify and estimate age, period and cohort components of change.

The significance of investigating these issues is twofold. First, this analysis will go a long way to explain whether the observed decline in marriage rates signify a real shift in marital behavior or, simply due to alterations in sampling design and question format (or (mis)understanding of the question by the respondent), associated with survey data. The results will expose pitfalls and highlight strengths in the different surveys on the variable, marital status, which is usually used as a control variable in most regression analyses. Repetition of identified sampling mistakes will thus be avoided in subsequent surveys.

Second, it is public knowledge that marriage has a direct and/or indirect relationship with other issues of concern, such as education attainment, labor force participation, fertility and race relations, among others. The essence of understanding the state of marriage, provided by the results of this study is to enhance development of policies in these areas which interlink with family policies. The need arises because of the negative relationship associated with most of the above-mentioned issues, especially for women. For instance, in the same period that marriages declined for women in South Africa, labor force participation increased, suggesting that marriage contributes to the price women pay in order to participate in the labor market. This stands as a concern since the institution of marriage is demonstrated to be important in both family and society in Africa. At the same time, narrowing gender gap in labor force participation is a longstanding objective in many economies.

We consider this study to be important. As mentioned earlier, most of the studies on marriage have been done elsewhere but Africa. This study will therefore contribute to knowledge in that regard. South Africa is an exceptionally interesting and applicable case

especially because the decline in marriage rates and other notable changes[6] (affecting marriage pattern) resulting due to the change in the socio-political climate are remarkable and also because of the tremendous restructuring of the marriage contract (for example the 1998 Recognition of Customary Marriage Act)[7] and the changing policies in labour market after apartheid which may actually aggravate these patterns. A clearer understanding of evidence underlying behavior in the marriage market is important in the evaluation of policies directly and indirectly affecting marital behaviour.

1.2 Research Objectives

Following the above background information, the main intent of the research is to understand the trends in the marriage patterns observed in South Africa. In particular, the study aims to establish the source of the change in marital behavior among African women, and consequently determine whether the change in marriage patterns signifies a permanent change in their marital behavior.

2.0 Literature Review

2.1 Review of Theoretical Literature

Becker's (1973) theory of marriage is pioneering work in the economics of marriage, providing a framework for analyzing behavior of issues concerning marriage using principles of economics. One of the implications of Becker's theory is that marriage is a manifestation of utility maximizing behavior. This is mainly because of its assumption that marriage is practically always voluntary, and hence it is expected to raise utility level above the alternative of remaining single. Two people will therefore agree to marry if the expected utility of being married exceeds that of being single. Central to Becker's theory is his argument that men and women bring different attributes to the marriage and each has specialized roles which account to the gains from marriage. This notion is known as positive assortative mating of complimentary traits. With reference to wages, for instance, as the wage differential between men and women narrows, presupposing increase in economic activity for women narrowing the economic and social gap for the two genders, women's incentive to marry decreases.

---

[6] These changes include increase in labour force participation, urbanisation, to mention but a few.
[7] The 1998 Recognition of Customary Marriage Act directs that entry into customary marriage must be with consent from both partners. In addition, the law regulates the dissolution of customary marriages and thus ensures the protection of women. This suggests that costs of marriage dissolution for a woman are lower than they were prior to the legislation.

Oppenheimer (1988) also suggests that as women's earnings increase, their desirability as marriage partners may extend across more years of their lives. According to Weiss (1997), this may result from the stochastic process associated with meetings between potential spouses. With each meeting, the partners compare their characteristics and evaluate their potential gains from marriage, while each of them anticipates own contribution to the marriage. If the gains from marriage exceed the expected gains from continuing with the searching for both, the partners marry; otherwise, they depart and wait for the next meeting to occur. This extended search time may allow women to be more selective of their marriage partners. Furthermore, women with extended time available for searching could afford to be most selective early in adulthood, when they can anticipate a large number of potential future partners. The implication is that longer search time results in later and fewer marriages, on average.

Becker's model also implies that marriage behavior is highly affected by the sex ratio, which proxies the marriage market. An increase in sex ratio, for instance, increases the demand for wives resulting in increasing marriage rates. Evidence in the literature supports the hypothesis that sex ratios influence the decisions to enter marriage. In considering the marriage market, issues of how to calculate sex ratio come in. This involves going beyond the numeric supply of potential partners (Becker, 1981 and Oppenheimer, 1988). Considerations of power relations between the two genders (Guttentag and Secord, 1983), marriageability (Wilson and Neckerman, 1986; Wilson, 1987), demography (for example age of marriageable partner) and geographical boundaries where meetings with potential partners are likely to occur are highlighted. Regardless of how sex ratio is defined, identical behaviours is predicted under similar conditions.

Wilson's (1987) argument that in availability of marriageable men results in less marriages for women has spurred fruitful research into changing marriage patterns and has mostly been applied to declining marriage rates among African Americans (for example, Wood 1995). Although the findings have been mixed, his arguments have generally found empirical support.

2.2 Review of Empirical Literature

As mentioned earlier, delay and decline in the proportions of marriage among young adults observed in most of the developed world has inspired Social Scientists to understand marital behaviour. However, most of the work on this issue has been done outside economics. Since the publication of Becker's (1973) theory, few economists have produced empirical analyses of marriage in the developing world. Using data from the early 1970s in the Philippines, Boulier and Rosenzweig (1984) provide confirmation of Becker's theory of marriage. They demonstrate that although the effect of education on age at marriage is exaggerated in models that treat education as exogenous, additional schooling leads women to marry later. Brien and Lillard (1994) show that when one controls for the effect of delayed marriage on education, that is, for the potential endogeneity of education, later age at marriage among women in Malaysia is explained in large part by increased enrollment and attainment. Consistent with Becker's prediction, with increased schooling the opportunity cost of marriage rises for women.

Equally consistent with Becker's theory predictions is the influence of sex ratio on marital behaviour. While a lot of empirical evidence[8] exists to support the demographic approach framework to show that marriage market conditions, such as quantity of men (sex ratios), cause decline in marriage, documentations in the literature[9] equally support the economic orientation framework, that is, quality of available partners explains this phenomenon. For instance, less educated men may not find employment and this may be less attractive to women as marriageable partners.

3.0 The Survey Data

We use synthetic panel data constructed from twelve successive cross sectional national surveys from the 1995-1999 October Household Surveys (OHSs) and the 2000-2006 Labour Force Surveys (LFSs)[10]. For each year of the survey data, our variable of interest is the question on marital status. The October Household Survey (OHS) has six categories namely: married (civil), married (traditional/customary), living together with partner, widower/widow, divorced/separated and never married and the categories

---

[8] Qian and Preston (1993) and Schoen and Wooldredge (1989) for example showed that some proportion of changes in first marriage rates in the United States could be explained by shortage of men.

[9] Fo example, Brien (1997) used five definitions of marriageable men: 1) all men; 2) all employed men; 3) all men who were employed, in school or short-term unemployed; 4) all men who were full-time employed; and 5) all men with earnings above a certain amount.

[10] Both the OHS and LFS were conducted by Statistics South Africa and a questionnaire was used for data collection.

represented in the Labour Force Surveys (LFS) include married or living together as husband and wife, widow/widower, divorced/separated and never married. Statistics South Africa (2001) defines marriage to include marriages by civil or religious ceremonies, marriages by customary law, and all those living together as married. In this regard, and for consistency with the classifications in the LFS, we will merge the first three categories in the OHS as married, and together with the widowed and the divorced, we categorise the group as "ever married". We will therefore estimate the proportion of a particular age cohort that is either ever married or single.

The OHSs are annual independent cross-sectional surveys, and different samples were designed for each of them. However, in a standard OHS, the sample was explicitly stratified by province, magisterial district, urban/rural and population group. A sample was drawn by applying a two-stage sampling procedure. In the first stage, enumeration areas (EAs) represented primary sampling units (PSUs) were systematically selected by means of probability proportional to size[11] principles in each stratum to ensure adequate representation. The measure of size was the number of households in each PSU. The database of EAs, as established during the demarcation phase of Census 96, constituted the sampling frame for selecting EAs. In the second stage, households were randomly selected from the selected EAs. The number of sampled households range from 16000 in 1996 through to 20000 in 1998 and 30000 in 1995, 1997 and 1999, depending on the availability of funding.

The LFS replaced the OHS after their discontinuation in 2000. The sampling design for the standard LFS is similar to that of the OHS. The main feature distinguishing these two surveys is that the LFS were designed to be a twice yearly rotating panel with the waves running in March and September. A twenty percent rotating scheme was designed, implying that new dwelling units would be included to replace the dropped ones in the second wave (SADA, 2001). For the purpose of this study, we will only make use of cross-sectional information and the September wave, in particular.

We restrict our sample to women between the ages 15 and 49, since this is the prime age range in which marriages take place.

---

[11] A disproportionately larger number of PSUs were allocated to the smaller provinces than the bigger provinces.

4.0 Methodology

4.1 Descriptive Approaches: The relationship between age and marriage in the cross-section and synthetic panel

As pointed out by Casale and Posel (2002), using independent cross-sections for comparative reasons is not without concern, mainly due to consistency issues, definition of variables, question coverage and sampling methodology. For the marriage study in particular, this is a concern because Statistics South Africa changes the way they categorize marital states. In other surveys, they separate marriage from cohabitation, while in others, they combine the two. This makes more appropriate studies of disaggregating marital options (married, single, and cohabitation) a far-reached cry for researchers.

In terms of sampling methodology, the 1996 and the 1998 surveys sampled only 16000 and 20000 households, respectively, relative to the "usual" 30000 in the other surveys, which raise concern over comparability of the results.

Another concern which applies to all surveys is little coverage of questions on marriage, in particular, and other questions generally useful in a marriage study. For instance, timing of marriage, which is important, for example, in the determination of timing of childbirth, and ultimately fertility, is missing. In addition, a highly culturally diverse society like South Africa would be expected to have highly culturally divided behaviors. Unfortunately, some of the surveys did not ask questions which would accommodate a culturally disaggregated study. At best, one would do a racial disaggregation, but this will not be adequate to establish behavior of a particular ethnicity[12], for example.

Nonetheless, amidst these and more data limitations, the study proceeds to give an aggregated picture of marriage trends, giving highlights of likely biases of the data, and attempting to explain them.

We will commence with a descriptive analysis where cross-sectional and cohort dimensions will be uncovered, while presenting the relationship between age and

---

[12] For example, within the indigenous African race, there are a number of ethnicities with varying cultures. Zulu culture may be different from a Xhosa culture regarding how marriages are traditionally handled.

marriage. Accordingly, we will estimate the proportions of a particular age cohort that falls into one of the four marital categories captured in the labour force and October Household surveys. To do this, we employ locally weighted estimation of scatter-plot smoothing (LOWESS) technique, which is explained in detail in appendix A. This technique originated from Cleveland (1979) and can be thought of as a series of nonparametric linear regressions at different points appropriately stitched together. The advantage of using nonparametric estimation is that it does not impose a functional form on the data but allows the data to determine the parameters and the shape of the curve (Deaton (1997). The method is fairly simple and yet yields fruitful "visual" results.

4.2 Decomposition Analysis: Age-Period-Cohort (APC) Model

An event such as marriage is influenced by age, period and cohort, in some way. Consequently, a more revealing and a deeper structure of underlying marital status schedules can only be extracted by disentangling age, period and cohort effects. The age-period-cohort model is the generational model workhorse used to identify whether changes in human behaviour are cohort-based or due to other factors such as age or calendar year. Age effect represents different allocations between the marital states associated with different age groups, giving a typical age profile. Period effect represents variation in distribution across marital states over time, simultaneously associated with all age groups, and affects every cohort. Cohort effect represents the difference in the distribution across groups of individuals with the same birth year, giving secular trends (Fu et al, 2004).

However, the well-known problem with specifying all three in an additive model is the identification problem resulting from multicollineality. Due to the exact linear dependency nature of the age-cohort-period model, i.e. $Age + Cohort = Period$, a unique solution for each of the effects cannot be found, since matrix containing age, period and cohort variables does not have full rank.

A number of solutions to this problem have been proposed in the literature (for example Glenn 1976; Fienberg and Mason 1978, 1985; Hobcraft, Menken and Preston 1982; Wilmoth 1990; Deaton 1997; O'Brien 2000; Fu 2000; Knight and Fu 2000; Yang, fu and Land 2004; McKenzie 2005; Yang 2006 and Yang and Land 2006). One of the simplest solutions is to drop one of the multicollinear variables. This approach, however, becomes

problematic where all of age, period and cohort effects are seen to be potentially important in explaining behaviour. A different proposal that acknowledges the importance of all three involves constructing a cohort that is represented by an age band to represent one cohort, so that the variables are no longer collinear.

Another approach involves normalizing age, period and cohort effects. A number of authors have proposed different normalizations. The constrained generalized linear modelling approach, also known as the coefficient-constraint approach, has been conventionally used among demographers and other social scientists, where constraints suiting the phenomenon under study have to be imposed. The problem with this methodological approach is that the parameter restrictions have to be theoretically motivated, which poses as a challenge for most phenomenon including the current study. In addition, the estimates are sensitive to the choice of restrictions.

Deaton's (1997) normalization assumes that period effects have zero mean and are orthogonal to the time trend. Thus, the trends observable to the data are attributed to age and cohort effects. Again, this approach is not immune to structural assumptions, but where no theory is available, this approach is useful as a descriptive device.

The intrinsic estimator approach utilizes estimable functions that are invariant to the selection of constraints on parameters (refer Fu, 2000). The basic idea of the intrinsic estimator is to remove the influence of the design matrix on coefficient estimates. Intrinsic estimators are also viewed as a special form of principal regression estimator (see Sen and Stivastava 1990, for the basics of principal components regression), that removes the influence of the null space of the design matrix on the estimator. It specifically estimates a constrained parameter vector b_o that is a linear function of the parameter vector b of the unconstrained APC accounting model. (Refer Yang, Schulhofer-Wohl, Fu and Land, 2007, for the algebraic, geometric and verbal description of the intrinsic estimator).

Yang, Land and Fu (2004) did a methodological comparison of the constrained generalized linear model and the intrinsic estimator methods using mortality data. From their results, they conclude that the intrinsic estimator appears to offer a useful

alternative to the conventional constrained generalised linear modelling approach in terms of the key statistical properties such as the finite-time-period bias and efficiency.

With no strong prior knowledge of the life-cycle theory of marriage, and since all these approaches have weaknesses, this study will use three approaches, as descriptive devices of the marriage trends. We will use Deaton's identifying assumptions. As a robustness check, will we use the intrinsic estimator approach, because of its efficiency properties and the age grouping approach, because it does not need any restrictions and hence its results are free from sensitivity of restrictions chosen.

4.3 Model Specification

In linear regression form, the APC model for marriage rates can be written as:

$$M_{ij} = \mu + \alpha_i A + \gamma_j P + \kappa_k C + \varepsilon_{ij} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \text{(1)}$$

Where M_{ij} denotes mean marriage rates for the *ith* age for $i = 15,...,49$ at the *jth* time period for $j = 1995,...,2006$; μ denotes the intercept; α_i denotes the *ith* row age effect; γ_j denotes the *jth* column period effect; κ_k denotes the *kth* diagonal cohort effect for $k = 1,...,35$ cohorts, with $k = i + j - 1995$; and ε_{ij} denotes the random errors with expectation $E(\varepsilon_{ij}) = 0$. A, P and C are matrices of age, period and cohort, respectively.

Due to the grouped nature of the data, the dependent variable is a proportion hence we need to perform a transformation on the observations. The problem is that the result from the transformation will be missing values, which will be subsequently dropped from the estimation sample. A better alternative, which we will adopt for our study, is to estimate a generalized linear model, GLM, instead of ordinary least squares, OLS, a method proposed by Papke and Wooldridge (1996). This is analogous to using logit and probit models in place of linear probability models, in order to ensure that the predicted probabilities are in the valid range of zero and one. By estimating a GLM, the predicted values of the fractional dependent variable will fall between the range of zero and one. We will accordingly estimate age, period and cohort effects by quasi-likelihood estimation methods which do not need any special data adjustments for the extreme values of zero and one (Papke and Wooldridge, 1996). The quasi-likelihood method is fully robust and

relatively efficient under the GLM assumption. In this regard, we maximise a Bernoulli log-likelihood function, given by:

$$l_i(\beta) \equiv M_i \log[G(x_i\beta)] + (1 - M_i)\log[1 - G(x_i\beta)] \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (2)$$

Where $G(.)$ is a logit link function, ensuring that predicted values of the dependent variable lies in $(0,1)$; β is a vector of parameters of age, period and cohort effects; and M_{ij} denotes marriage rates.

Following the preceding discussion, we will specify a generalised linear model with a logit link and a binomial family formulated as:

$$\theta_{ij} = \log\left(\frac{m_{ij}}{1 - m_{ij}}\right) = \mu + \alpha_i A + \beta_j P + \gamma_k C \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (3)$$

Where θ_{ij} is the log-odds of marriage and m_{ij} is the probability of marriage in cell (i, j). We will include the robust option in the GLM to obtain robust standard errors which will be particularly useful if we have misspecified the distribution family.

5.0 Empirical Findings

5.1 Marriage Trends in South Africa: 1995-2006

For starters, this proved to be a fruitful exercise as far as checking the quality of the data is concerned. In tracking marital states across time, I noticed a major inconsistency in the results from the 1998 OHS. It emerged that none of the kids who had not yet reached the legal age to marry, were single. And it turned out that they were all divorced. Similarly, quite a high proportion of people over the age of 80, still have their spouse around. Clearly, the code for "married" was used for "never married"; "never married" for "divorced"; "divorced" for "widowed"; and "widowed" for "married". It turned out that the inconsistency was produced by incorrect documentation that had been sent out with the survey. The 1997 questionnaire had been distributed with a 1998 cover sheet. The marital codes had changed in the 1998 questionnaire and this led to erroneous coding of the categories. These are clearly critical errors and especially because marital status is one variable that is included in most regression analyses.

### 5.1.1 Stylized Facts from Cross-Sectional Analysis

This section presents the empirical findings of the research. The stylised facts are presented for both cross-sectional and cohort data analyses examined in a univariate

context, by age. We first present findings from the cross-sectional analysis, giving visual evidence of the change in trends in marriage pattern for the period 1995 to 2006. The aim is to provide a picture of how the marriage rates have changed over the years for African women. These trends are compared across gender and racial lines. This is followed by evidence of the change in trends by cohort analysis, demonstrating that the observed change in the cross-sectional analysis may be a generational change, suggesting possible permanent change in marital behaviour among African women in the period under study. Finally, we cement the analysis by presenting results from the APC analysis, determining whether the change in marital pattern is influenced by age, period or cohort effects.

Each trajectory in the cross-sectional graphs shows proportions of a particular marital state for a particular age in a particular year. The balance at each age for each year for the married state is single, widowed or divorced, implying that a decline in married status is an increase in either or all of the other states.

Figure 1: Cross-Section Analysis: Proportion Married

Source: Own Calculations Using OHS 1995-1999 and LFS 2000-2006 (September Waves)
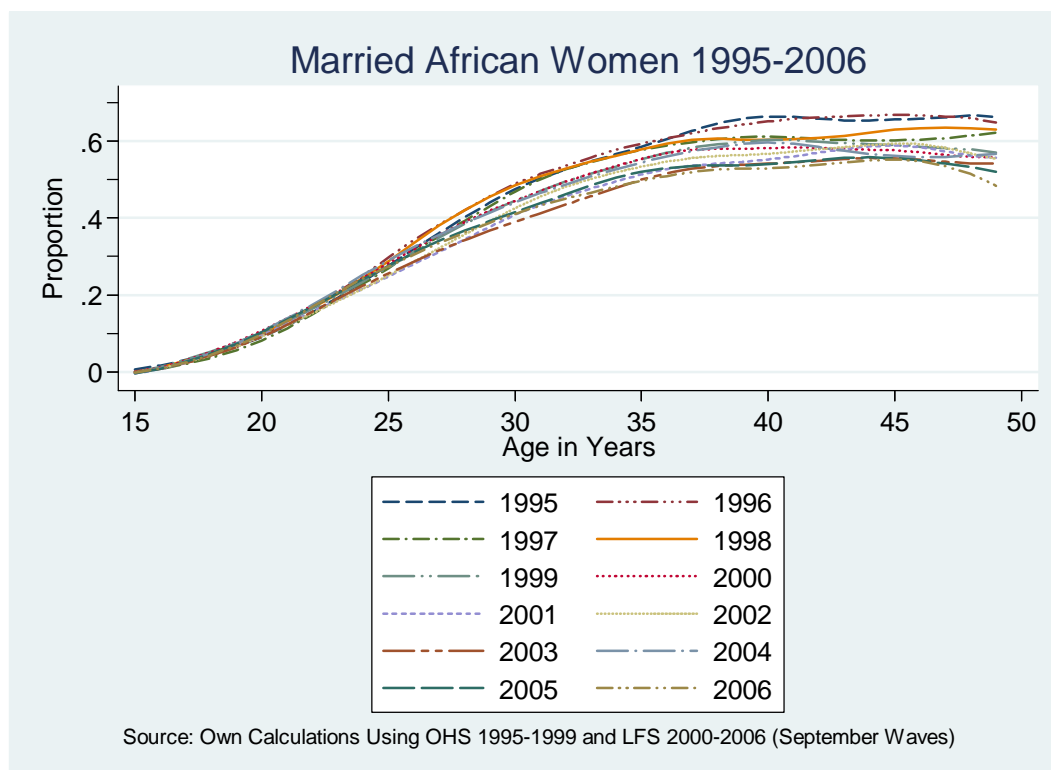
From the diagram of married state in figure 1 above, we observe clear evidence of varying marriage rates over the period 1995-2006. The cross-sections provide a snapshot

pointing towards a decline in marriages. The decline is evident in all the ages beyond 25. The diagram for the single state (in appendix B) adequately mirrors the one for married state, for all ages, with the deviations captured in the widowed and the divorced states (also in appendix B). This suggests that singlehood in the form of not entering marriage is what adequately accounts for the decline in marriage, as opposed to singlehood resulting from marriage dissolution by death of a spouse or divorce.

For each cross-section, proportion that is married increases with age, and rapidly so, until the age of 35, after which it remains fairly stable. This suggests that when African women get to the mid-thirties, they are unlikely to enter into marriage. We also note that for each cross-section, a significant proportion of African women start getting married after the age of 22, giving evidence of late entry into marriage. Calculations of the singulate mean age at marriage[13] for all the years is not less than 29, indicating evidence of late marriages. This conclusion draws from a proposition made by Bogue (1969), that the median age of 22 marks the benchmark for late entry into marriage for females. However, for the most part, we observe low levels of married proportions in the younger ages because they are still in school.

Closely looking at the diagrams for the widowed state, we note almost zero activity in the ages ranging from 15 to 35. Beyond the age of 35, some variation in the widowed state is evident across the cross-sections. Even though the direction of the variation is unclear, the picture generally shows that widowhood increases with age. However, the levels are generally low. At the younger ages, widowhood might be explained by high mortality rates among men, compared to women, with deaths resulting from violent acts.

Similarly, the picture from the divorced state shows that divorce rate increases with age, and the levels are low across the cross-sections. Even though the picture is pointing towards a decline, we cannot deduce whether there is an overall decline in divorce, since the variation in the cross-sections is not systematic. In general, since divorce rates are minute in the ages that marriages are highly likely to dissolve, it appears that African women are reluctant to enter into marriage but once they get married, they stay married. Thus, death seems to be the main channel of marriage dissolution.

---

[13]

Behaviour for the 1996 and 1998 trajectories in the married state, which coincide, stand out for ages between 25 and 35 where the proportions are consistently above all other trajectories. One possible explanation for this behaviour is that, for unknown reasons, Statistics South Africa sampled more married African women in this age range, in these years. The 2003 trajectory behave in an opposite manner, in the same age range, suggesting that less African married women were sampled.

The cross-section analysis for African men gives a similar story[14]. From the trajectories (in appendix C), we observe similar pattern to that of women, supplying emphasis of evidence of the decline in marriages among African women.

Marriages among White women has also declined (See picture in appendix D). However, compared to their African women counterparts, this decline is not as systematic and it especially concentrates in the 20-30 age range, rather than in the whole age distribution of 15-49. For that reason, White women may not necessarily be forsaking marriage indefinitely. The age range where the declined is observed for White women suggests that they may be postponing marriage to attain additional education.

In general, as depicted in Table 1, the percentage decline of marriages is greater for males than it is for females, with the decline greatest among African men at 16 percentage points, followed by White men. Among women, the decline is also larger for African women at about 10 percentage points. Interesting, though, is that while marriages seem to be going on a downhill for everybody else, Indians experience an increase in marriages of 7 percent. This increase may be explained by the enactment of the Recognition of Customary Marriages Act in 1998, with results of increased number of marriages registered.

Table 1: Mean Marriages and Standard Deviations by Gender and Race, 1995 and 2006)

	Male		Female		% Change	
	1995	2006	1995	2006	Male	Female

---

[14] We considered African men to be the potential partners for African women since marriage across racial lines rarely occur in South Africa, following the legacy of apartheid. For example, the 1949 Mixed Marriages Act prohibited marriage between persons of different population groups. The 1950 Immorality Act also banned sexual relations between Whites and non-Whites and the 1950 Area Act made it compulsory for people to live in an area designated for their classification group.

All races	.3741	.3155	.3992	.3596	−15.7 −9.9
	(.4839)	(.4647)	(.4897)	(.4799)	
Africans	.3231	.2717	.3496	.3149	−15.9 −9.9
	(.4677)	(.4448)	(.4768)	(.4645)	
Coloureds	.423	.4004	.4233	.4168	−5.3 −1.5
	(.4941)	(.4901)	(.4941)	(.4931)	
Indians	.5349	.5298	.5748	.6158	−0.1 7.1
	(.499)	(.4996)	(.4946)	(.4868)	
Whites	.6248	.5444	.6719	.6297	−12.9 −6.5
	(.4842)	(.4982)	(.4696)	(.4831)	

Standard deviations in parentheses

However, aware that it is widely accepted that the OHS and the LFS surveys cannot be easily compared due to different sampling designs of different cross-sections, we closely look at the cross-sections to establish if we can notice major differences in marriage rates between them. We do this for only females of all race groups. From Table 2, it is worth noting that there are no dramatic variations in behaviour between and within the two surveys, enhancing comparability. For instance, we observe stability in marriage proportions when OHSs get replaced by LFSs in 2000, and less variation between successive cross-sections.

Table 2: Average Marriage rates in South Africa: 1995-2006

GENDER = Female

| **Race** | **Year** | | | | | | | | | | | |
	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
African	.35	.36	.34	.35	.33	.33	.32	.33	.31	.34	.32	.31
Coloured	.42	.45	.42	.43	.43	.45	.42	.44	.44	.43	.42	.42
Indian	.57	.56	.56	.58	.55	.6	.52	.53	.58	.61	.61	.62
White	.67	.61	.64	.62	.61	.65	.65	.64	.62	.67	.61	.63

The Cohabitation Factor

One of the limitations of this study is failure to net out the extent of cohabitation from the married state due to inconsistencies in the categorization of the marital states for the different surveys. For instance, out of the twelve surveys, married and cohabitating states were lumped together from the years 2000 to 2003. Failure to disentangle levels of cohabitation from the married state would give a misleading impression in the dynamic analysis of marriage patterns, especially where cohabitation behaviour is similar to that in the developed world. In that regard, high levels of marriages would be inflated by cohabitation, and the decline in marriages may be understated. This kind of data capturing thus becomes a limitation if one wants to investigate whether marriages are being replaced by single-hood or cohabitation.

We show, in Table 3 below, the extent of this problem, by factoring out cohabitation from the married state in the surveys where the two were categorized separately. Married proportions are clearly inflated when cohabiters and the married are lumped together, as evidenced by the figures in the years 2000 to 2003 in Table 2 above. Cohabitation over the twelve-year period has increased by 7 percentage points. This suggests that African women's union formation behaviour may be similar to that of the western world, with informal union replacing the formal union. However, whether cohabiting relationships in South Africa tend to be long-term or short-term is another issue which needs further investigation.

Table 3: Extent of Cohabitation among African Women

|	1995	1996	1997	1998	1999	2004	2005	2006
MARRIED	.3057	.3167	.2886	.2746	.2675	.2318	.2105	.2036
	(.4607)	(.4652)	(.4531)	(.4463)	(.4427)	(.422)	(.4077)	(.4027)
COHABITING	.0439	.0477	.0539	.0726	.0649	.106	.1095	.1113
	(.2049)	(.2131)	(.2258)	(.2595)	(.2464)	(.3079)	(.3122)	(.3145)

Standard deviations in parentheses

The evidence from the cross-sectional analysis suggests that marriages are postponed, not only among African women, but across the gender and racial lines as well. However, in contrast to White women, the postponement appears to be indefinite for African women, as suggested by marriages profiles across the age distribution. In the next section, we explore, by use of cohort data, the hypothesis that the decline in marriages of African women is real.

### 5.1.2 Stylised Facts from Cohort Analysis

Cross-sectional analysis may give an impression that there has been a change in marital behaviour, which may not entirely be reliable since individuals are aging throughout the cross-sections. We therefore strengthen our findings from the cross-sectional analysis by undertaking a cohort analysis and this aids us to assess the developments in the marriage market in the first decade of democracy. Specifically for the current study, cohort analysis enables the study to determine whether the changes observed in marriage rates are generational changes. Cohort analysis is more informative than just comparing subjects of the same age in different years. Consequently, rather than comparing an age cohort in different cross-sections, we reroute to following an age cohort overtime.

Cohort : Ever Married

Age of cohort defined by age in 1995
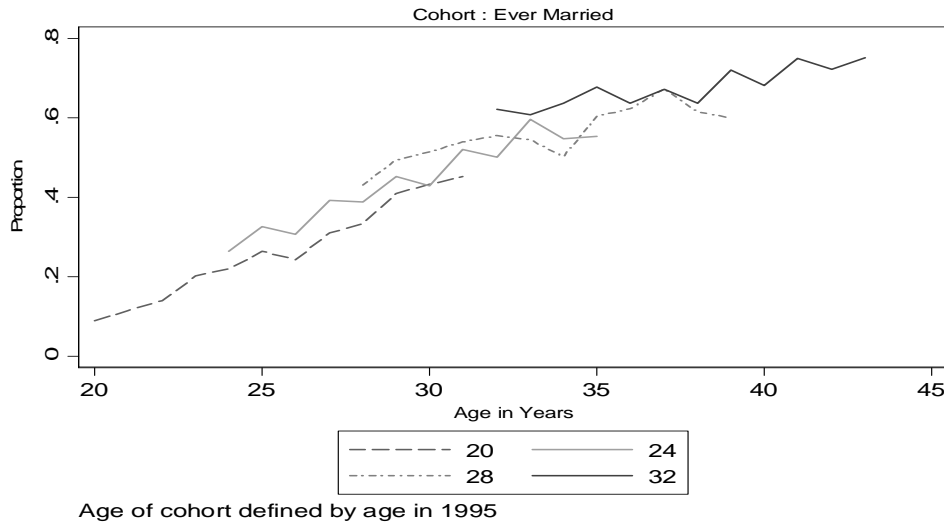
The picture from the cohort analysis in figure 2 above shows that within cohort, there is a positive relationship between age and proportion that is married. Between cohorts, we observe that for the most part of the period span, trajectories of older age cohorts lie on top of the ones for the younger cohorts, suggesting different marital experiences for different generations. For instance, comparing 20 year old age cohort at the age of 25 and 24 year old age cohort at the same age, we observe a clear shift in marital behaviour, with the proportion married for the younger cohort, lower than their predecessors. Similar comparisons can be made between different other cohorts, with results pointing to a decline in marriages.

Figure 3: Cohort Analysis for five-year age bands
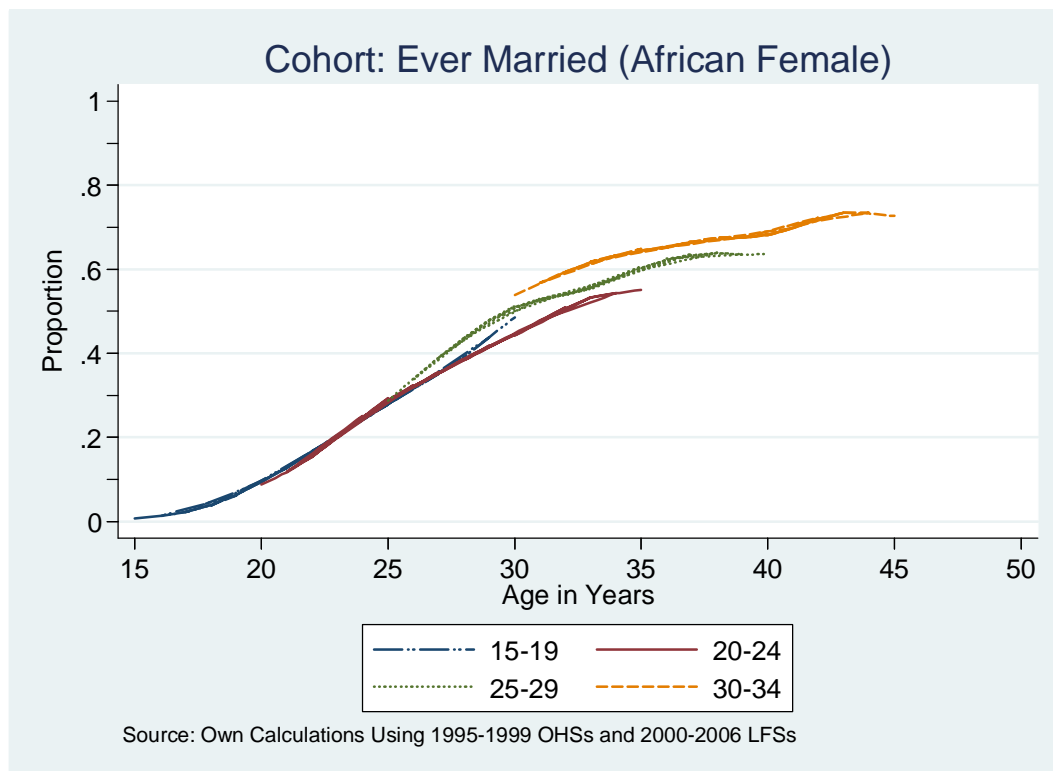
The trajectories are not smooth, demonstrating cohort-global effects. With survey data, concern is more of sampling effects rather than macroeconomic effects. By grouping cohorts in five-year band, to smoothen the trajectories, shown in figure 3 above, we obtain clear evidence of differences in marital behaviour between the young and the old. At the age of 35, 50 percent of the cohort aged 20-24 is married, while at the same age, about 65 percent of the cohort aged 30-34 is married. We therefore conclude from the cohort analysis that within cohorts, proportion that is married increases with age, and between cohorts, marital experiences are different, signifying a generational change. However, we do not rule out presence of sampling effects, following the zigzagged trajectories in figure 2.

The picture from the cohort analysis contains age, cohort and period effects. The age and cohort effects are clearly seen. For the cohort effects for example, the trajectories for the older cohorts are always on top of the younger cohorts (in the five-year age band except for the 15-19 and 20-24, which coincide). This suggests that the young and the old face different marital experiences, signifying a generational change.

Age effect is also evident by the positive relationship between age and proportion married and takes the "slight" S shape. In this study, it is evident that proportion that is married tends to increase rapidly between the ages 25 and 30, and less rapidly beyond 30. Though the age effect is discontinued for the later ages in this study, proportions that is married is expected to be fairly constant in the ages beyond 50 when no or less marriage formations take place, already formed marriages get stable and deaths get stable too. The proportion is expected to start falling in old ages when marriage dissolutions, usually by death, get little or no replacement by new entrants into the marriage institutions.

The period effects are captured by the span of each trajectory, which corresponds to the same 12 years of the surveys, 1995 to 2006. Graphs (in figure 4 below) from APC models provide disentangled age, period and cohort effects.

5.2 Decomposition Analysis: Age, Period and Cohort Effects

Following the decomposition strategy discussed in the methodological section, we present the estimation results, separating the age, cohort and period effects in the marriage rates. But since no theoretically motivated restrictions were made, we use the APC model as a descriptive device, which is adequate for our study objectives.

To enhance reliability of our results, we employ three different procedures (discussed in the methodology section) for identification of the APC model. The top row of the figure contains the age, period and cohort effects from Deaton's identifying assumption. The middle row displays the same, but for the three-year age groupings in a cohort and finally, the bottom row captures estimates from the intrinsic estimator.

Figure 4: Age, Period and Cohort Effects

Decompostion Analysis: APC Models

Source: Own Calculations using 1995-2006 Synthetic Panel
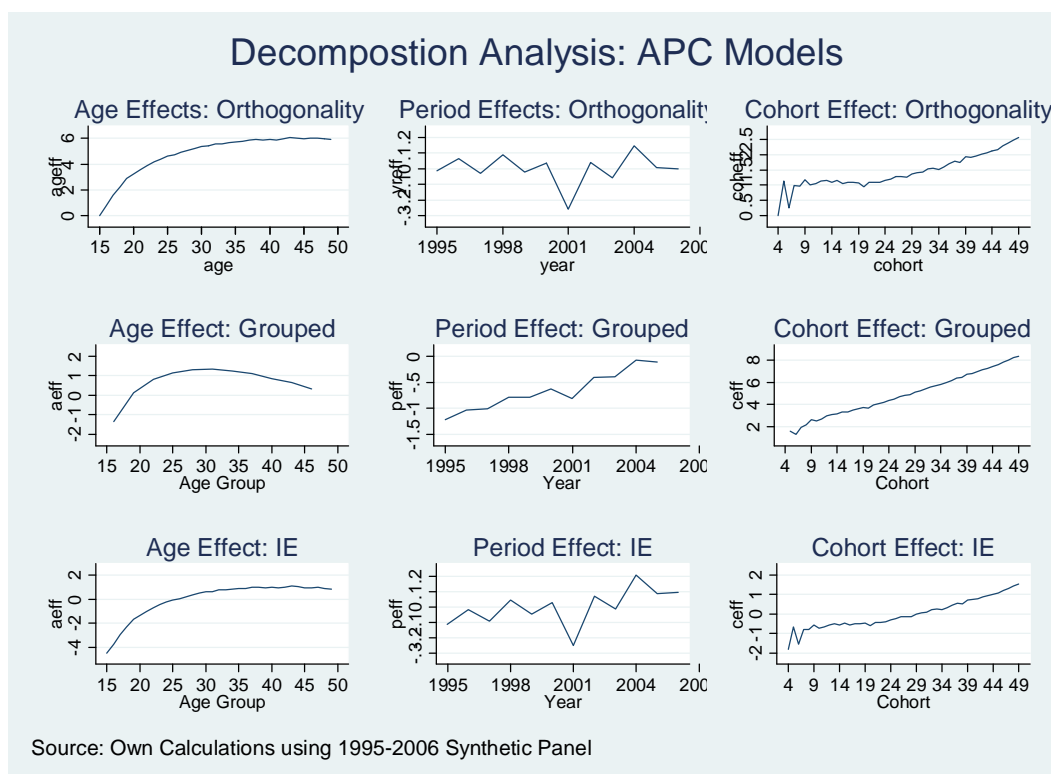
From all three identifying assumptions, the age effect is positive for all cohorts, but increases with age. The young, by and large, show low positive values. This in part, may be because they are still in school. A rather steep slope is observed up to the age of 30, a pick-period for marriages. Beyond the age of 30, age effect remains fairly constant or decline (for the three-year grouping normalization). Perhaps behaviour after the age of 30 is due to women's economic independence, having established themselves in the workplace. The age effect, on its own, provides undisputable evidence of decline in marriage. With the age profile remaining near constant (or declining) after the age of 30, it means that proportions married between ages 31 and 49 will be low. The low contribution of the 31-49 age range to the proportion married will result in low total marriages observed for the whole age range of prime marriages, resulting in observed decline in marriages. More evidence of the decline in marriages is revealed by the cohort effect coefficients.

Just like the age effect, the cohort effect is also consistently positive for all age cohorts. Very low and constant effects are present for the young generations, up to the age of 24, after which increasing effects mount. The implication is that cohort effects are present

22

and the two generations behave differently, driven by their own cohort influences. Again, the immediate explanation of this is that the young are still in school, while the older are ready for marriage. The association between cohort and marriage is characterised by a gradual and persistent decline in marriage rates in successive cohorts. In similar lines to the discussion of the age effect above, each cohort contributes less to the marriage proportion, resulting in observed decline in total marriages.

The year on year differential of the period effect is negligible, except for the more apparent 2001. This suggests that 2001 may be an outlier as far as sampling of women is concerned, and should not be compared with the other surveys. Mostly period effect revolves around zero. This is a very important result because in enhances our confidence to conclude that the decline in marriages observed in the cross-section and cohort analysis is a real artefact, which is not just driven by the sampling and questionnaire design. This decline in marriages reflects a push for African women to achieve economic, educational and career goals, enhanced by the opportunities enabled by the new legislature.

What is even more important is that the evidence the decline in marriages generated from these surveys can be relied upon since the results are not sensitive to changing identifying assumptions. Basically, we get the same shapes and story even though the magnitudes of the effects differ for the different identifying assumptions. This makes our results valid and reliable, despite the data concerns. In the next section, we explore, in a univariate context, some of the likely reasons for the decline in marriages.

5.3 Likely Explanations of the Decline in Marriages

Overall, the results obtained from the cross-section, cohort and decomposition analysis give evidence of a real decline in marriages. However, an assessment of the results obtained in this study proves to be taxing since studies addressing similar questions about marriages in South Africa or elsewhere could not be found. Nevertheless, we rely on theory of marriage and studies on other trends in South Africa to discuss the likely explanations of decline in marriages.

a) **Women's Economic Status**

Studying if migration patterns have changed in the post-apartheid South Africa, Posel (2003) identified two changes at the same time as female labour migration increased. First, marriage rates among African females (15 years and older) fell from 35 percent in 1993 to 30 percent in 1999. Second, household composition changed in that the average number of employed men resident in rural African households decreased, and the proportion of households with no employed male members increased, suggesting an increase in female-headed households. From her findings, Posel (2003) concluded that the increase in female labour migration may be consistent with the maximising decisions of households in which women are less likely to enter formal unions. The implication is that in doing so, women face less male resistance to migrate. This is consistent with Becker's predictions that if the perceived advantages of marriage are outweighed by its disadvantages, there will be no commitments. Evidence of increase in female labour force participation in South Africa is adequately documented (for example Casale and Posel (2002), who used data from the 1995-1999 OHSs to study the continued feminisation of the labour force in South Africa).
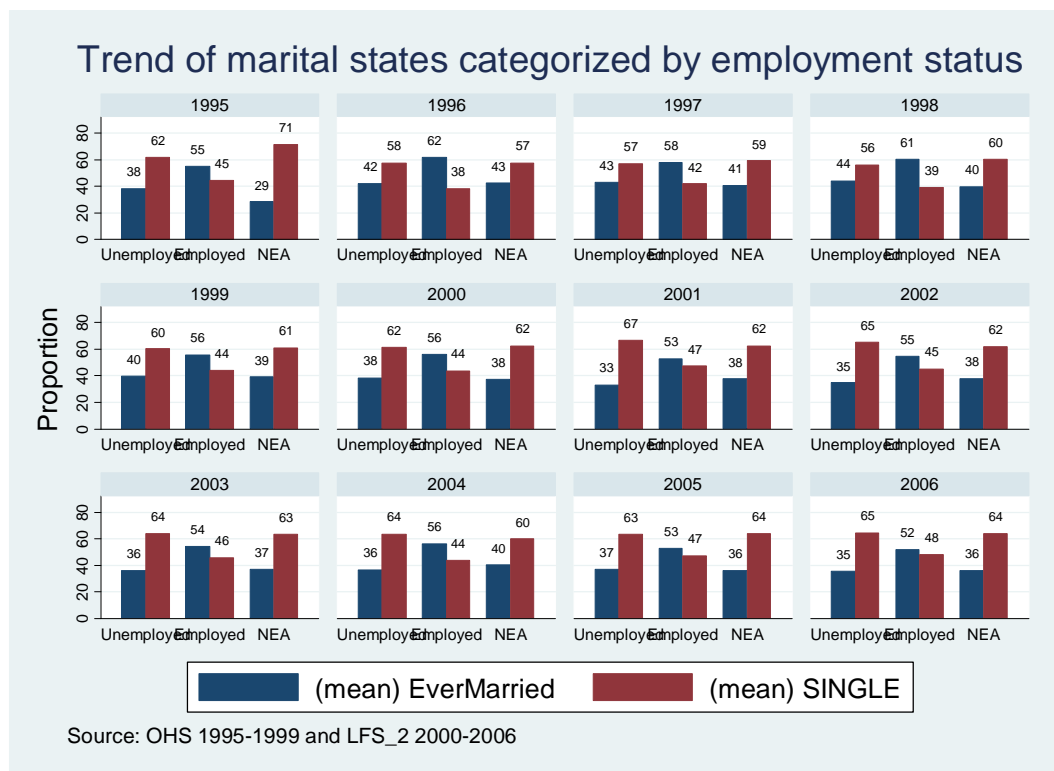
In this regard, the question that immediately follows is whether the lower proportions of women that are married corresponds to increase in economic independence, proxied by increase in labour force participation and education attainment. This question is answered by the trends below:

Trend of marital states categorized by education level

Source: OHS 1995-1999 and LFS_2 2000-2006
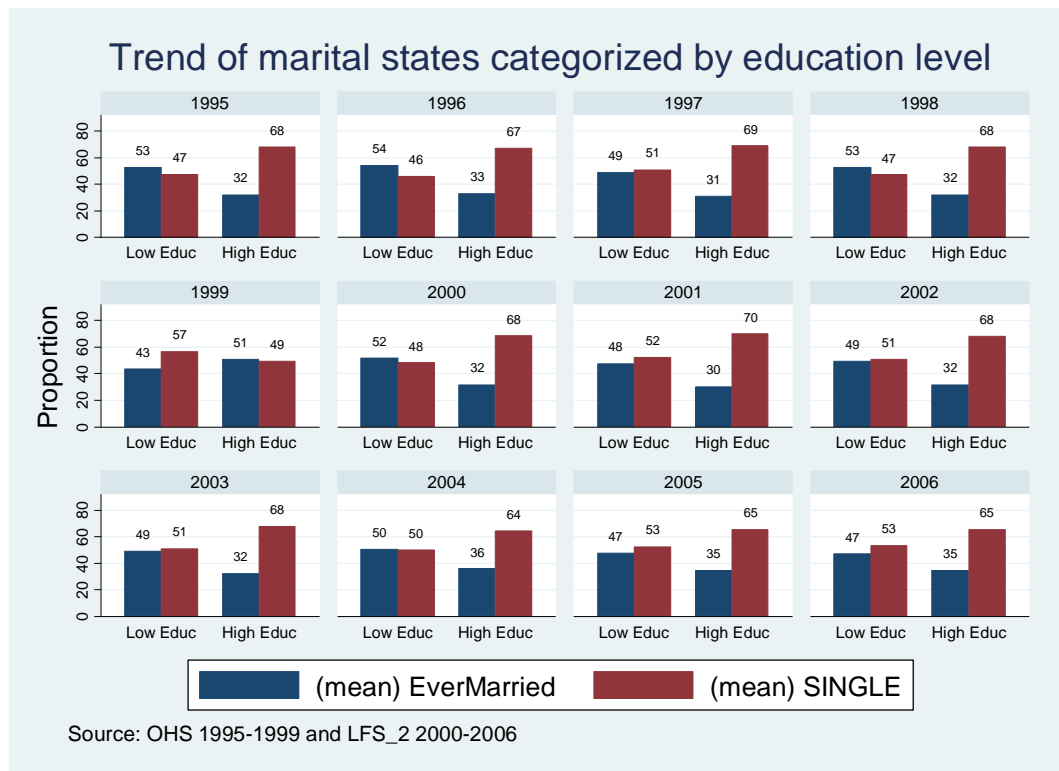
From this picture, we establish that, indeed, education plays a role in the marital state of African women. Among the low educated, the difference is almost negligible for the ever-married and single states. On the other hand, most of the highly educated women have never been married. This is consistent with Ntuli's (2007) findings. With regard to education in her investigation of determinants of female labour force participation between 1995 and 2004, Ntuli (2007) finds that secondary education slightly reduces the prospects of getting married. She however finds the lower and higher categories to be insignificant. The picture displayed for employment status, however, is contrary to the prediction of women's economic independence hypothesis. On the overall, we notice that the proportion that has never married is either not economically active or unemployed. The highest proportion of women that is married falls in the employed status. This is contrary to findings by Ntuli (2007) who finds that marriage considerably reduces the probability of women to participate in the labour force. She finds marriage as culprit to lower levels and extent of women's labour market participation. A possible explanation for our result is that quality of wife is an issue for South African men and they are possibly being picky when they are choosing spouses.

It is also possible that those who are employed are in inferior and low-paying jobs (after all, evidence (for example Casale and Posel, 2002) shows that feminisation of the labour market in South Africa has resulted into either unemployment or employment in inferior

and low-paying jobs), and if "selectiveness" of African men in the marriage market is anything to go by, then the this result may be unsurprising. Even those women who are not economically active may not be marriage materials for men, and that is why we observe such high proportions of never married women in that labour market category.


Trend of marital states categorized by employment status

Source: OHS 1995-1999 and LFS_2 2000-2006

Clearly, an investigation of what determines the choice of marriage decision in a multivariate context is worth undertaking in order to get a consistent and reliable explanations regarding why marriages have declined over the years. From her investigation of determinants of female labour force participation between 1995 and 2004, Ntuli (2007) finds that marriage considerably reduces the probability of women to participate in the labour force. She finds marriage as culprit to lower levels and extent of women's labour market participation. With regard to education, Ntuli (2007) finds that secondary education slightly reduces the prospects of getting married. She however finds the lower and higher categories to be insignificant. In sum, marriage tends to hinder participation in the labour force, and hence translates into postponing and/or forgoing marriages in order to enhance economic activities.

b) **Sex Ratios**

Sex ratios, as implicated by Becker's marriage theory may also help explain this marriage behaviour. Increase in sex ratio portrays a favourable marriage market. Our descriptive analysis of sex ratio pattern over the years shed some light in connection with marriage behaviour in South Africa.

Figure 5: Sex Ratios

Graphs by YEAR
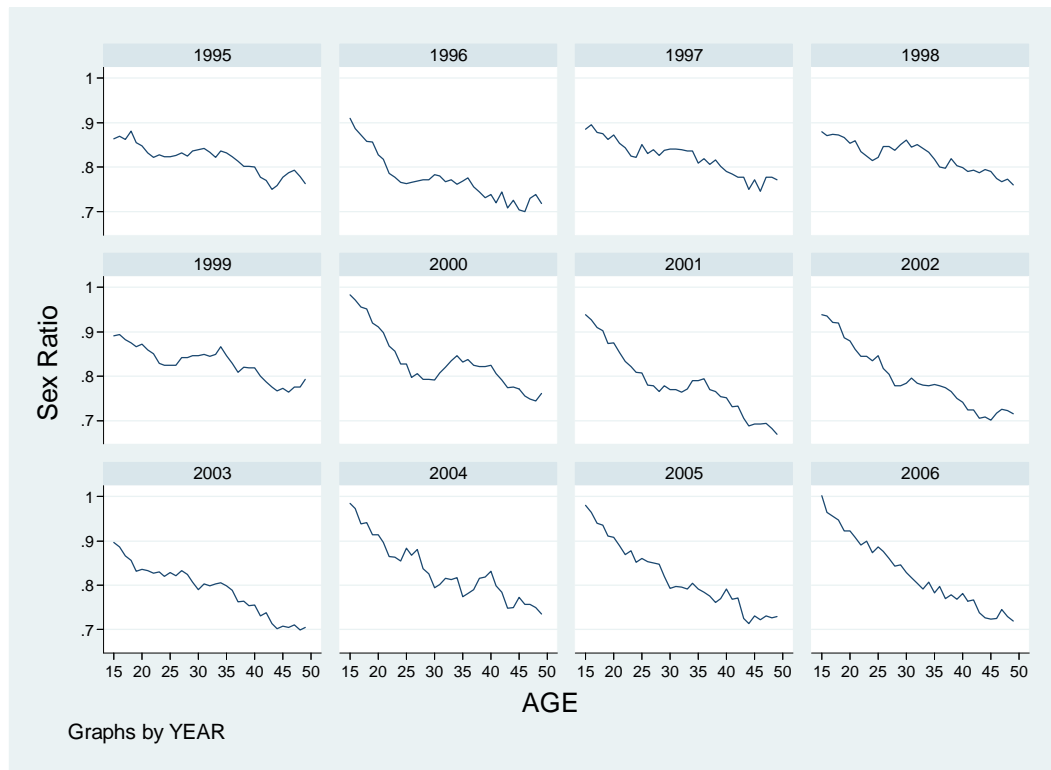
With sex ratio over one signifying more men than women, and hence increasing demand for wives, we note, from figure 5, that throughout the period under study, sex ratios are below one, for all the ages. In particular, sex ratio are even much lower for those ages where marriages are highly likely than when the young are still in school. One of the explanations for low sex ratios is high mortality among men, compared to women, due to unnatural causes resulting from violent behaviour. Shortage of men may thus explain why marriages are low in South Africa. Related to this is poor marriage market, as measured by shortage of quality marriageable men. With the narrowing gender gaps as far as economic status is concerned, women become picky when searching for potential mates. And with high levels of unemployment in South Africa, less and less marriages prevail.

In the next chapter, we hypothesize that marriage and labour supply decisions are jointly determined. This endeavour is motivated by a trend similar to the one for marriage

observed in the labour market. Labour force participation for African South African women increased, alongside the decline in marriage (Casale and Posel, 2002). The objective of that chapter will be to establish the reasons why marriages have declined, in a multivariate context.

In the third paper, we will model assortative mating, by education, which might give an insight of the situation in marriage market. This will be insightful and establish if women (and/or men) are becoming picky, and hence, justify the decline in marriages. Ultimately, we believe that this research will help further understanding of this important institution and accordingly, guide public policy.

6.0 Conclusion

We set out in this study to examine marital behaviour in South Africa between 1995 and 2006. We used a synthetic panel which was constructed from the October Household Surveys and the Labour Force Surveys. Cohorts, which were defined by year of birth, were followed through the period. Our main objective was to net out cohort effects from the age and period effects in the marriage rates in order to establish if marital experiences of younger and older cohorts are different. Although age effects were present and bigger in magnitude, most importantly, cohort effects were also evident, indicating a shift in marital behaviour between the younger and the older cohorts, pointing to decline in marriage. We identify greater economic independence and shortage of marriageable men as likely explanations for the decline.

While sceptics may argue that decline in marriage on its own may not be interesting, we feel that it may be of public concern when certain public policies either intentionally or unintentionally affect the decision to marry. For instance, it is widely known that marital status, fertility and labour supply are jointly determined. Thus, shifts in marital behaviour may shift behaviour in childbearing and labour supply decisions.

7.0 Appendices

Appendix A: Descriptive Approaches: LOWESS Technique

Basically, LOWESS is a method of smoothing scatter-plots in which the fitted value at x_k is the value of a line fit to the data using weighted least squares where the weight for

(x_i, y_i) is large if x_i is close to x_k and small if x_i is further away from x_k. Plotting the smoothed points, which form a nonparametric regression of Y on X, frequently allows the perception of effects on the scatter-plot that are otherwise difficult to detect. However, caution must be taken so that the scatter-plots must not be over/under-smoothed. A tradeoff between bias and variance arises with respect to degree of smoothing such that over-smoothing increases bias but reduces variance while under-smoothing does the reverse. One simple solution towards optimal smoothing is by choosing an optimal window, within which a point can be smoothed. Within the window defined by the bandwidth, weighted least squares regression of y on x is run using the following weights:

$$w_i = \left(1 - |u_i|^3\right)^3$$

Where $u_i = \dfrac{x_k - x_i}{d_i}$ and d_i is the distance between x_k and its furthest neighbour within the band. The weighting allocates a value of 1 for $x_k = x_i$ and declining weights for points further away, reaching zero at the borders of the bands. Thus, points near the middle of the window receive more weight than points towards the boundaries.

For this study, a constant bandwidth of 0.3 is chosen (following Wittenberg, 2002). This indicates that 30 percent of the sample is contained in each band and is used to smooth each point. LOWESS technique thus fits a line to data, portion by portion. The fitted x_i are used as $f(x_i)$ in the model $Y_i = f(x_i) + \varepsilon_i$ where ε_i is assumed to be a white noise process. The result of this process is that the predicted smoothed Y values for n observations is obtained by 'stitching' together n regressions (Hamilton (2004). In this case, $f_j(x_i)$ is the proportion in each marital state k; $k = married, \sin gle, widowed, divorced$ controlling for age, gender and population group. We will therefore have a nonparametric estimation which will create a scatter of points which follow a distinct trend of plots of proportions on the y-axis of each age categories on the x-axis that is in a particular marital status. Such representations will be drawn for the years 1995 to 2006. The slopes of the resulting graphs from the analysis give the estimates of the underlying flow rates thereby giving information about the changing levels in the marital categories, over the years.

Specifically interested in levels and flows, this graphical analysis will identify the key turning points and other broad trends in the data. The following functions $f_1 \ldots f_4$ will be estimated nonparametrically using LOWESS technique:

$$P_{ji} = f_1\left(age_i, popgrp_i, gender_i \mid x_i\right)$$

Where P denotes proportions in a particular marital state; subscript j is as previously defined; and x_i is a vector of controlling variables. The method is fairly simple and yet yields fruitful results.

Appendix B: Single, widowed and divorced states of African women

Figure 6: Proportion single for African women

Source: Own Calculations Using OHS 1995-1999 and LFS 2000-2006 (September Waves)
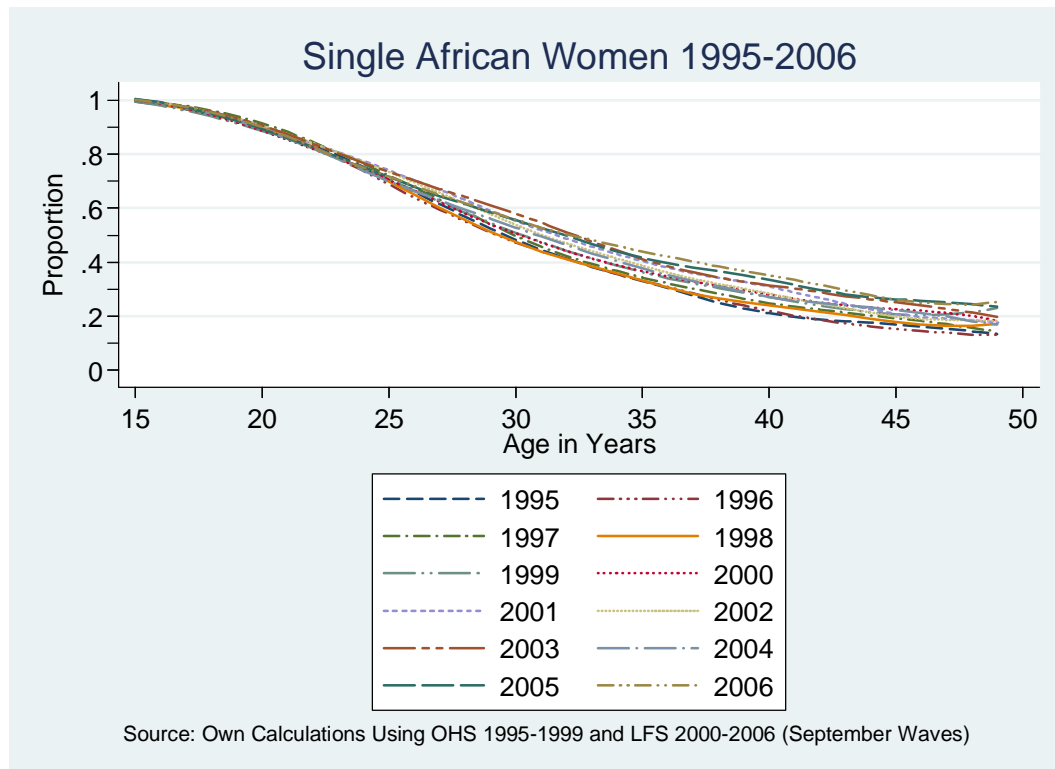
Figure 7: proportion widowed for African women

Widowed African Women 1995-2006

Source: Own Calculations Using OHS 1995-1999 and LFS 2000-2006 (September Waves)
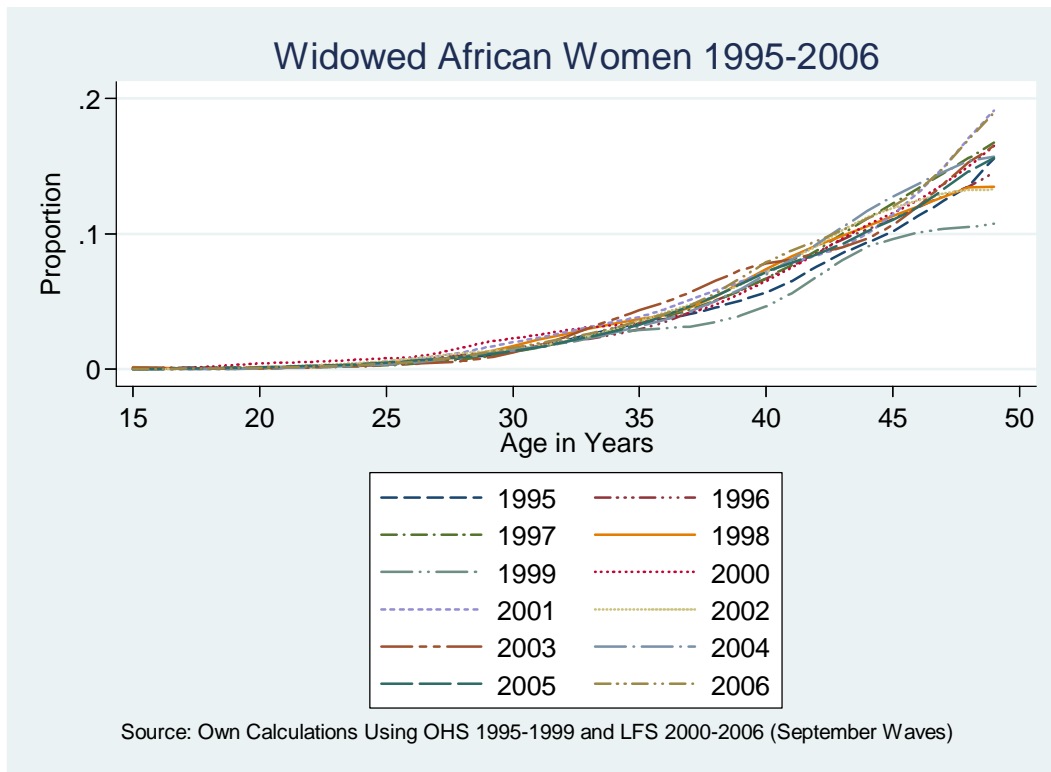
Figure 8: Proportion divorced for African women

Divorced African Women 1995-2006

Source: Own Calculations Using OHS 1995-1999 and LFS 2000-2006 (September Waves)
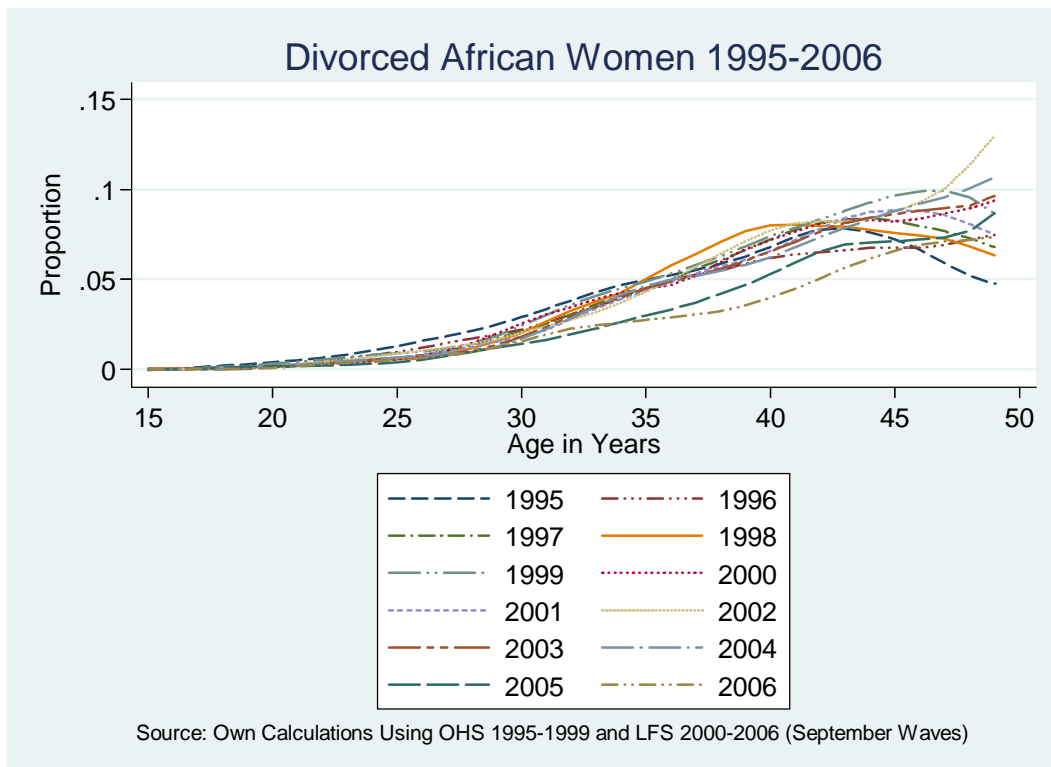
Appendix C: Marital states for African men

Married African Men 1995-2006

Source: Own Calculations Using OHS 1995-1999 and LFS 2000-2006 (September Waves)
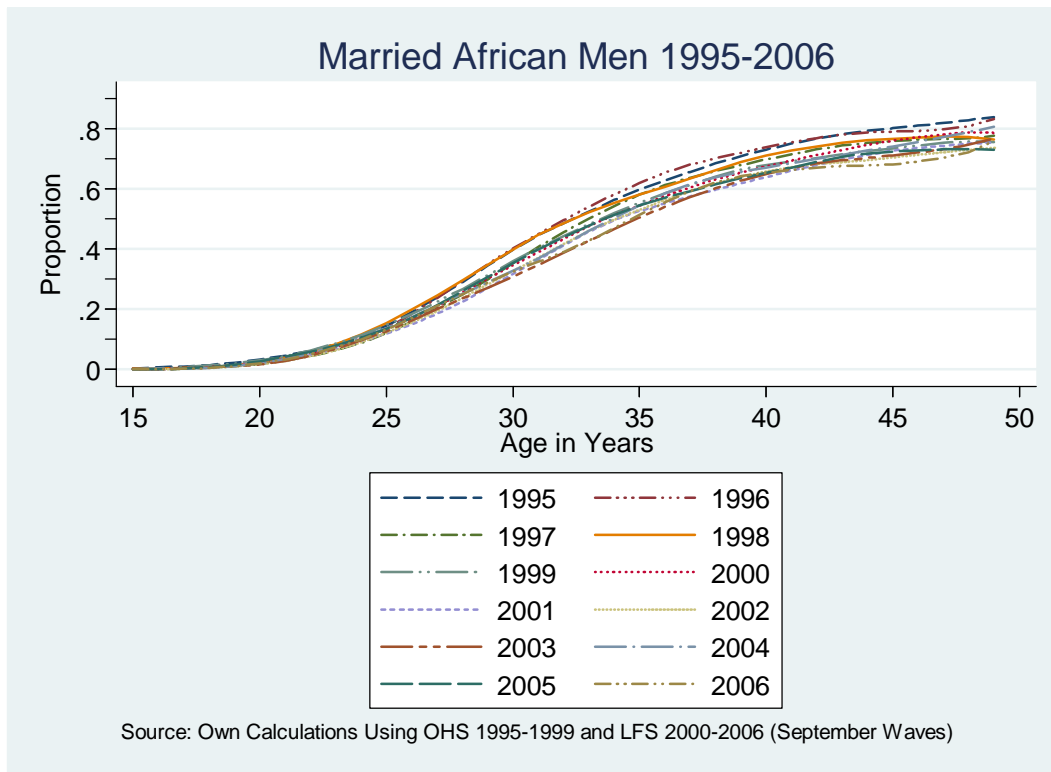
Figure 10: proportion single for African men

Single African Men 1995-2006

Source: Own Calculations Using OHS 1995-1999 and LFS 2000-2006 (September Waves)

Widowed African Men 1995-2006

Source: Own Calculations Using OHS 1995-1999 and LFS 2000-2006 (September Waves)

Figure 12: proportion divorced for African men

Divorced African Men 1995-2006

Source: Own Calculations Using OHS 1995-1999 and LFS 2000-2006 (September Waves)
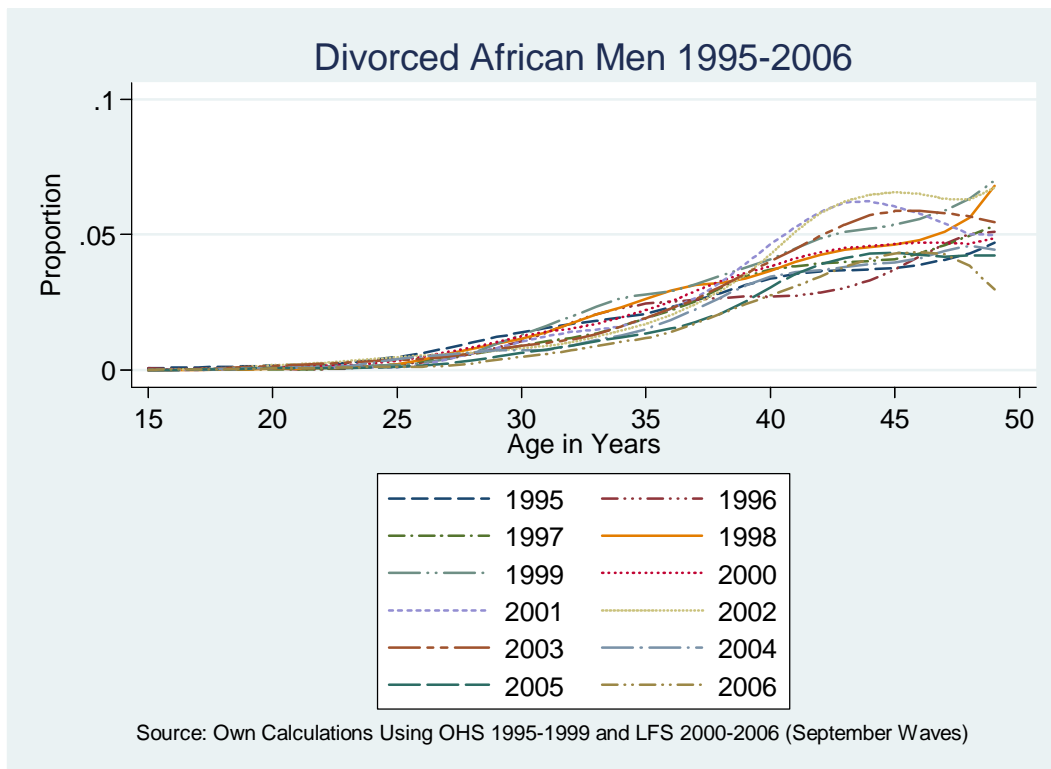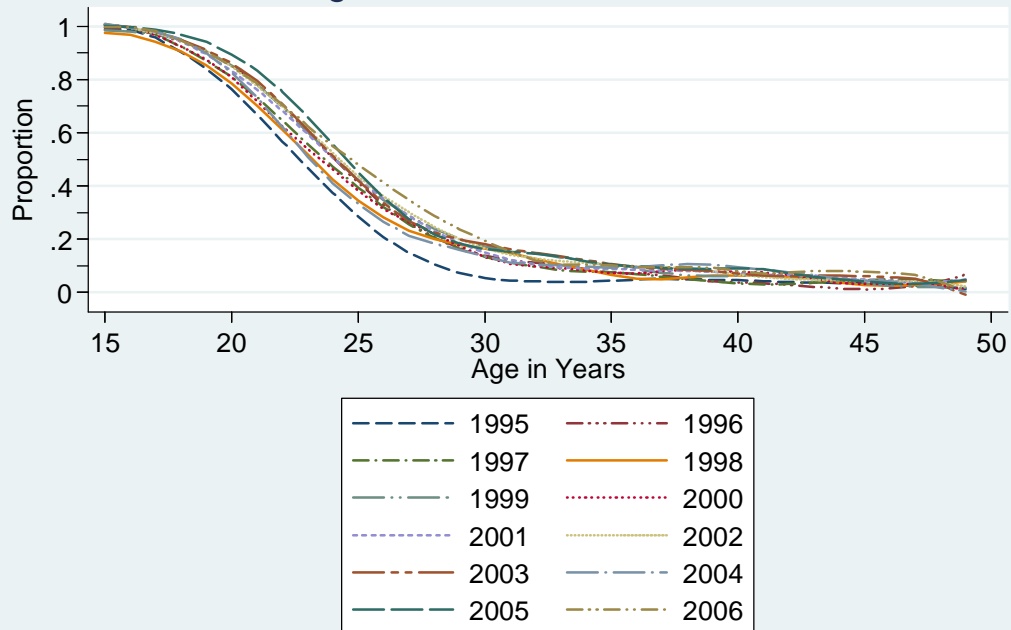
Appendix D: Marital states for White Women
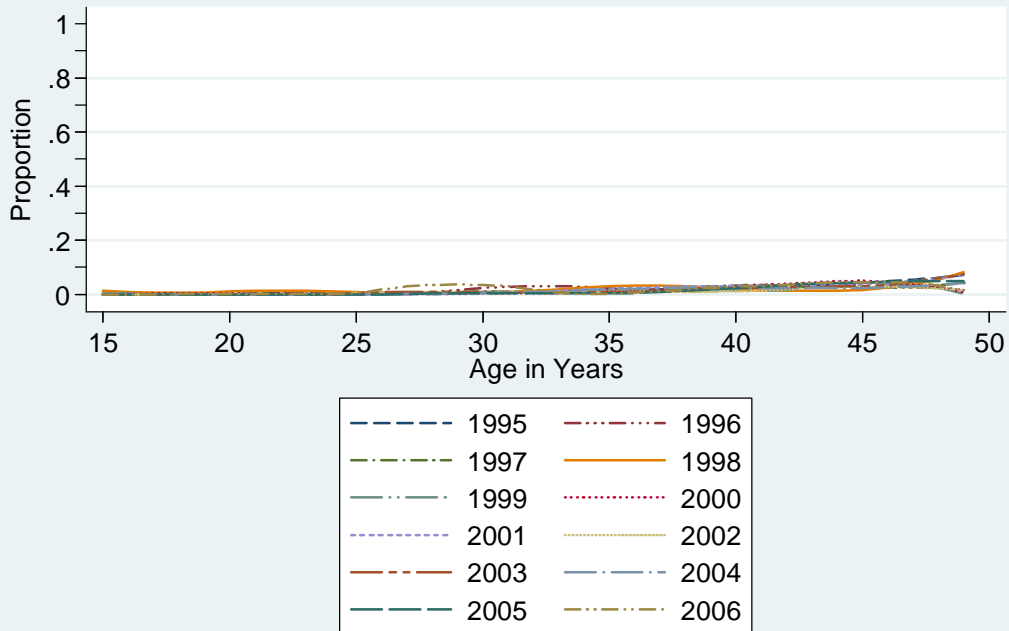
Married White Women 1995-2006
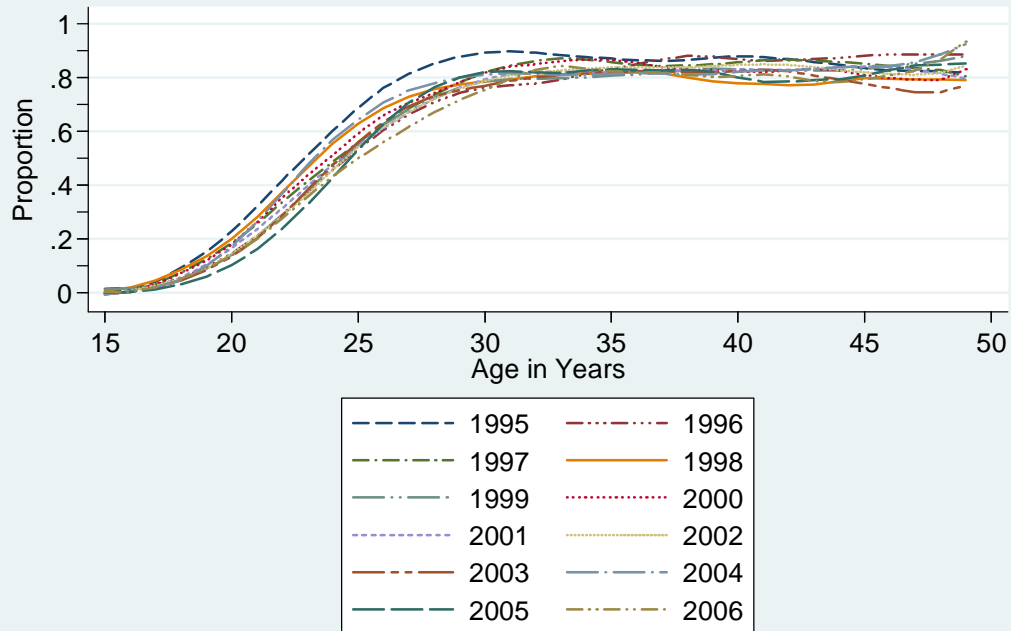
Source: Own Calculations Using OHS 1995-1999 and LFS 2000-2006 (September Waves)
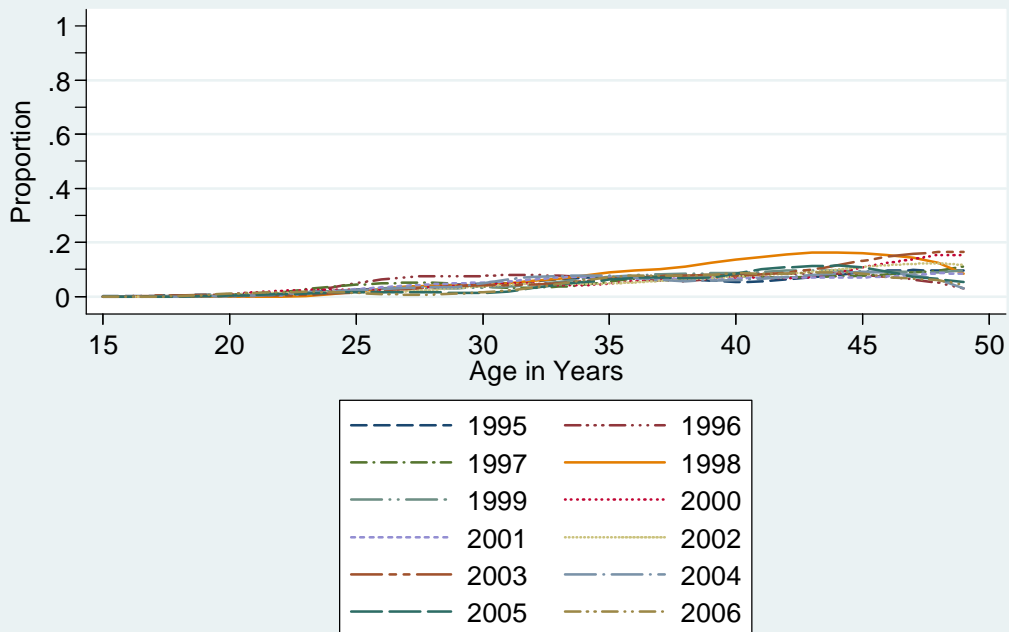

Single White Women 1995-2006

Source: Own Calculations Using OHS 1995-1999 and LFS 2000-2006 (September Waves)

34

Widowed White Women 1995-2006

Source: Own Calculations Using OHS 1995-1999 and LFS 2000-2006 (September Waves)

Divorced White Women 1995-2006

Source: Own Calculations Using OHS 1995-1999 and LFS 2000-2006 (September Waves)

References

Adegboyega, O., (1994), The situation of families in East and Southern Africa, in Kalule-Sabiti, I., Palamuleni, M., Monde Makiwane, M and Amoateng, A. Y., (2007), Family formation and dissolution patterns in Families and Households in the Post-apartheid South Africa, Amoateng, A. Y. and Heaton, T. (Eds), *HSRC Press*, Cape Town.

Becker, G.S., (1981), A Treatise on the Family, Cambridge, *Harvard University Press*.

Becker, G.S., (1973), A Theory of marriage: Part I, *The journal of Political Economy*, Vol.81, No. 4, pp. 813-846.

Boulier, B.L. and Rosenzweig, M.R. (1984), Schooling, search, and spouse selection: Testing economic theories of marriage and household behaviour, *The Journal of Political Economy,* 92(4), 712–732.

Bogue, D.J., (1969) Marriage and marriage dissolution, in Kalule-Sabiti, I., Palamuleni, M., Monde Makiwane, M and Amoateng, A. Y., (2007), Family formation and dissolution patterns in Families and Households in the Post-apartheid South Africa, Amoateng, A. Y. and Heaton, T. (Eds), *HSRC Press*, Cape Town.

Branson , N. and Wittenberg, M., (2007), The Measurement of Employment Status in South Africa using Cohort Analysis, 1994-2004, *South African Journal of Economics*, Vol 75:2.

Brien, M.J., (1997), Racial Differences in Marriage and the Role of Marriage Markets, *The journal of Human Resources,* Vol. 32, No.4, pp741-78

Brien, M.J., and Lillard, L.A., (1994), Education, Marriage and First Conception in Malaysia*, The journal of Human Resources*, Vol. 29, No.4, Special Issue: The Family and Intergenerational Relations, pp1167-1204.

Casale, D. and Posel, D. (2002) "The Continued Feminisation of the Labour Force in South Africa: An Analysis of Recent Data and Trends", *South African Journal of Economics*, Vol.70 (1): 156-184.

Cleveland, W., (1979), Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association,* Vol. 74, pp829-836.

Deaton, A., (1985), Panel Data from Time Series of Cross-sections, Journal of Econometrics, 30, 109-126.

Deaton, A., (1997), The Analysis of Household Surveys: A Microeconometric Approach to Development Policy, *The Johns Hopkins University Press*, Baltimore.

Fienberg, S.E., and Mason, W.M., 1978, Identification and Estimation of Age-Period-Cohort Models in the Analysis of Discrete Archival Data, *Sociological Methodology*, Vol. 8, Pp. 1-67.

Fienberg, S.E., and Mason, W.M., 1985, Specification and Implementation of Age, Period, and Cohort Models, *Cohort Analysis in Social Research*, Pp. 45-88.

Fu, W.J., 2000, Ridge Estimator in Singular Design with Application to Age-Period-Cohort Analysis of Disease Rates, *Communications in Statistics-Theory and Method* 29:263-78.

Fu, W.J., Hall P. and Rohan T.E., 2004, Age-Period-Cohort Analysis: Structure of Estimators, Estimability, Sensitivity, and Asymptotics, *Journal of the American Statistical Association,* revised and resubmitted.

Glenn, N.D., 1976, Cohort Analysts' Futile Quest: Statistical Attempts to Separate Age, Period, and Cohort Effects, *American Sociological Review*, 41:900-904.

Hobcraft, J., Menken, J. and Preston, S., 1982, Age, Period and Cohort Effects in Demography: A Review, *Population Index* 48:4-43.

Guttentag, M. and Secord, P.F., (1983), Too Many Women? Beverly Hills, Califonia: *Sage*.

Kalule-Sabiti, I., Palamuleni, M., Monde Makiwane, M and Amoateng, A. Y., (2007), Family formation and dissolution patterns in Families and Households in the Post-apartheid South Africa, Amoateng, A. Y. and Heaton, T. (Eds), *HSRC Press*, Cape Town.

Knight, K. and Fu, W.J., 2000, Asymptotics for Lasso-Type Estimations, Annals *of' Statistics* 28: 1356-78.

Ntuli, M., (2007), Exploring the Status of African Women in the South African Labour Market, 1995-2004, PhD Thesis, School of Economics, University of Cape Town, South Africa, Unpublished.

O'Brien, R.M., (2000), Age-Period-Cohort Characteristic Models, *Social Science Research,* Vol. 29, pp123-39.

Oppenheimer, V. K., (1988). A Theory of Marriage Timing, *American Journal of Sociology* Vol. 94 pp 563-91.

Papke, L.E. and Wooldridge, J.M., (1996), Econometric Methods for Fractional Response Variables With an Application to 401 (K) Plan Participation Rates, Journal of Applied Econometrics, Vol.11, No.6, pp619-632.

Posel, D., (2003), Have Migration Patterns in post-Apartheid South Africa Changed? *African Migration in Comparative Perspective*, Johannesburg.

Qian, Z. and Preston, S.H., (1993), Changes in American Marriage, 1972 to 1987: Availability and Forces of Attraction by Age and Education, *American Sociological Review*, Vol. 58, pp 482-495.

SADA (2001), Statistics South Africa: October Household Survey, 1999, SADA: South African Data Archive, Pretoria, http://www.nrf.ac.za.

Schoen, R. and Wooldredge, J., (1989), Marriage Choices in North Carolina and Virginia, 1969-71 and 1979-81, *Journal of Marriage and the Family*, Vol.51, pp465-81.

Sen, A., and Srivastava, M., (1990), Regression Analysis: Theory, Methods, and Applications, New York: Springer-Verlag.

Weiss, Y., (1997), The Formation and Dissolution of Families: Why Marry? Who Marries Whom? And What Happens Upon Divorce, In Handbook of Population and Family Economics, edited by Mark R. Rosenzweig and Oded Stark, Vol. 1A, pp81–124. Amsterdam: Elsevier
.

Wilmoth, J.R., (1990), Variation in Vital Rates by Age, Period, and Cohort, *Sociological Methodology*, Vol. 20, Pp. 295-335.

Wilson, W.J., (1987), The Truly Disadvantaged, *University Press*, Chicago.

Wilson, W. J., and Neckerman, K.M., (1986), Poverty and Family Structure: The Widening Gap between Evidence and Public Policy Issues, pp 232-59 in Fighting Poverty, edited by Danziger, S.H. and Weinberg, D.H. Cambridge, Massachussets: *Harvard University Press.*

Wittenberg, M., (2002), Job Search in South Africa: A Nonparametric Analysis, *The South African Journal if Economics*, Vol. 70:8.

Wood, R.G., (1995), Marriage Rates and Marriageable Men: A Test of Wilson's Hypothesis, *Journal of Human Resources*, Vol. 30, No. 1.

Yang, Y., Fu, W.J. and Land, K.C, (2004), A Methodological Comparison of Age-Period-Cohort Models: An Intrinsic Estimator and Conventional Generalised Linear Models, *Sociological Methodology*, Vol.34, pp75-110.

Yang, Y., (2006), Bayesian Inference for Hierarchical Age-Period-Cohort Models of Repeated Cross-Section Data, *Sociological Methodology*, Vol. 36 pp39-74.

Yang, Y. and Land, K.C., (2006), A Mixed Models Approach to the Age-Period-Cohort Analysis of Repeated Cross-Section Surveys, with an Application to Data on Trends in Verbal Test Scores, *Sociological Methodology*, Vol. 36 pp75-97.

Yang, Y., Schulhofer-Wohl, S., Fu, W.J. and Land, K.C., (2007), The Intrinsic Estimator for Age-Period-Cohort Analysis: What it is and how to use it, Revised and Resubmitted to *American Journal of Sociology.*

About DatatFirst

DataFirst is a research unit at the University of Cape Town engaged in promoting the long term preservation and reuse of data from African Socioeconomic surveys. This includes:

• the development and use of appropriate software for data curation to support the use of data for purposes beyond those of initial survey projects
• liaison with data producers - governments and research institutions - for the provision of data for reanalysis
• research to improve the quality of African survey data
• training of African data managers for better data curation on the continent
• training of data users to advance quantitative skills in the region.

The above strategies support a well-resourced research-policy interface in South Africa, where data reuse by policy analysts in academia serves to refine inputs to government planning.

Data**First**