UNIVERSITY OF SOUTHAMPTON

United Kingdom

The age difference between spouses: Cross-national and within-country variations

By Catherine Heather Mercer

Thesis submitted for the degree of Doctor of Philosophy

Department of Social Statistics

Faculty of Social Sciences

May 2002

THE AGE DIFFERENCE BETWEEN SPOUSES: CROSS-NATIONAL AND WITHIN-COUNTRY VARIATIONS

ABSTRACT

This thesis examines the spousal age difference, which is defined as the male age at marriage minus the female age at marriage. Considerable variation is identified both between and within countries in the average age difference, as well as in its constituent parts. In addition, a greater association is observed cross-nationally between the age difference and the age at which females rather than males marry, while within countries, there is evidence of the opposite pattern.

Several hypotheses have been proposed in the literature in an attempt to explain the variations in the age difference. However, these hypotheses have seldom been examined. This thesis examines three of the most common, theoretically complementary hypotheses. These are (1) that variations in the age difference reflect the age-sex structure of men and women eligible to marry; (2) the age difference is a by-product of factors associated with the timing of marriage; and (3) that the age difference reflects the status of women, which is associated with the idea that there are preferences regarding the age difference.

The thesis begins by using aggregate-level data from the United Nations' WISTAT3 database to consider cross-national variations in the average age difference. The hypothesis that these variations reflect the status of women is then examined by considering associations between the age difference and hypothesised indicators of women's status. Given the continuing debate surrounding how women's status is conceptualised and measured, this analysis interprets women's status in several different ways. Some evidence in support of this hypothesis is observed, but not when considered in light of the association between female age at marriage and, in turn, the age difference and women's status.

The thesis proceeds to use individual level data collected over the last decade by the Demographic and Health Survey for 30 developing countries. The analysis takes a lead from the frequently cited study by Casterline, Williams and Macdonald (1986) that examined variations in the age difference by using data from 28 World Fertility Surveys, which were conducted in the 1970s. Given the literature concerning the quality of data collected by surveys in developing countries, this thesis also examines the extent of age misreporting among the 30 samples. It is recognised that polygamous unions tend to have larger age differences, and since polygamy tends to be more prevalent in some regions than others, comparisons are also made with these data specifically for monogamous unions. Joint distribution theory is used to identify evidence in support of the hypothesis that the observed distributions of age differences are as expected, given the age-structure of those marrying. Since the hypotheses considered are complementary, the 'by-product hypothesis' is also examined. Graphical modelling is used to consider the association between the age at marriage of husbands and wives net of both marginal and conditional associations with a number of factors that are hypothesised as being associated with age at marriage. Since a direct association remains between the husband's age and wife's age net of these factors, it is concluded that the age difference is not simply a by-product of such factors.

The final section of the thesis uses data from the 1995 Egyptian Demographic and Health Survey to examine the women's status hypothesis at the individual level. Egypt is focussed on as its survey included questions about women's autonomy within the household. Evidence in support of this hypothesis is identified between the age difference and only a few of the indicators of women's autonomy. However, when graphical modelling is used to consider these associations net of the association between the age difference and age at marriage, there is no net association between the age difference and women's status. There is some evidence of this association, however, when the model includes other factors associated with age at marriage.

This thesis concludes that variations in the age difference tend to reflect the age-sex structure of men and women marrying. However, it is unlikely that these variations are also a by-product of factors associated with marriage timing. In addition, this thesis finds little evidence to support the popular women's status hypothesis, especially when considered net of the association between the age difference and women's age at marriage.

Contents

Contents (continued)

Contents (continued)

Contents (continued)

Contents (continued)

Contents (continued)

List of tables

List of tables (continued)

List of tables (continued)

List of figures

List of figures (continued)

List of figures (continued)

List of figures (continued)

Abbreviations (alphabetical)

ALG	Algeria
ARG	Argentina
AUL	Australia
AUS	Austria
BAH	Bahrain
BAN	Bangladesh
BEG	Belgium
BEN	Benin
BRA	Brazil
BUL	Bulgaria
BUF	Burkina Faso
CAM	Cameroon
CAN	Canada
CAR	Central African Republic
CHL	Chile
CHN	China
CMC	Century-Month-Code
COL	Colombia
COM	Comoros
CON	Congo
COS	Costa Rica
COT	Côte d'Ivoire
CUB	Cuba
CZF	Czechoslovakia (Former)
DEN	Denmark
DOM	Dominican Republic
DHS	Demographic and Health Survey
ECU	Ecuador
EGY	Egypt
ETH	Ethiopia
FIJ	Fiji
FIN	Finland
GHA	Ghana
GRC	Greece
GUB	Guinea-Bissau
GUT	Guatemala
GUY	Guyana
HAI	Haiti
HKO	Hong Kong
HUN	Hungary
IND	India
INS	Indonesia
IRE	Ireland
IRQ	Iraq
ISR	Israel
ITA	Italy
MIM	Mixed Interaction Modelling
MLI	Mali
MOR	Morocco
MOZ	Mozambique
MYA	Myanmar
NEP	Nepal
NET	Netherlands
NIG	Niger
NOR	Norway
NZE	New Zealand
OMA	Oman
PAK	Pakistan
PAN	Panama
PAR	Paraguay
PER	Peru
PHI	Philippines
PNG	Papua New Guinea
POL	Poland
POR	Portugal
PUE	Puerto Rico
REU	Reunion
RWA	Rwanda
SAF	South Africa
SAR	Syrian Arab Republic
SAU	Saudi Arabia
SEN	Senegal
SIE	Sierra Leone
SIN	Singapore
SMAM	Singulate mean age at marriage
SOM	Somalia
SPA	Spain
SRI	Sri Lanka
SWA	Swaziland
SWE	Sweden
SWI	Switzerland
TAN	Tanzania
THA	Thailand
TRT	Trinidad and Tobago
TUR	Turkey
UAE	United Arab Emirates
UGA	Uganda
UKI	United Kingdom
UN	United Nations
URU	Uruguay
US	United States of America
USR	USSR (Former)
UZB	Uzbekistan

Abbreviations (alphabetical, continued)

JAM	Jamaica
JAP	Japan
JOR	Jordan
KAZ	Kazakhstan
KEN	Kenya
KRE	Korea (Republic of)
KUW	Kuwait
MAA	Mauritania
MAL	Malaysia
MAS	Mauritius
MAW	Malawi
MEX	Mexico
VEN	Venezuela
VIE	Viet Nam
WFS	World Fertility Survey
WISTAT	Women's Indicators and Statistics database
YUS	Yugoslavia (Former)
ZAI	Zaire
ZAM	Zambia
ZIM	Zimbabwe

ACKNOWLEDGEMENTS

Many people have helped me to start, continue, and complete this thesis.

I owe thanks to Dr Maíre Ní Bhrolcháin who has been my supervisor for my M.Sc. and Ph.D. theses. She encouraged me to study for a Ph.D. while completing my M.Sc., and provided the inspiration for this thesis. She enabled me to obtain an ESRC studentship (award number R00429734507), which provided funding for the first two years of study. I also thank her for the useful comments she has made on earlier drafts.

Dr Mac McDonald has also supervised this thesis and provided useful comments on earlier drafts, especially with regard to the statistical elements. In addition, I am grateful to Dr Dave Holmes who gave advice regarding data manipulation, and Dr Peter Smith and Fatima Salgueiro who gave advice on graphical modelling and using MIM3.1.

I am also grateful to the heads of the Department of Social Statistics, Professors Chris Skinner and Ray Chambers, for their advice and support throughout my Ph.D., especially since taking-up full-time employment at the Royal Free and University College Medical School, London.

Professor Anne Johnson, head of the Department of Primary Care and Population Sciences at the Royal Free and University College Medical School, London, has been very generous with study leave over the past two years. Equally important has been her encouragement and inspirational positive attitude. In addition, Dr Andrew Copas, Sandy Gale, Stephanie Black, and Patricia Kingori have been supportive in so many ways.

My family and friends have been amazing throughout my studies in terms of their continued encouragement, patience, and love. Above all, I thank my husband, Dr Paul Mercer, for his unwavering faith in me over the years.

Chapter 1:
Introduction & literature review

1.1 Introduction

The significance of marriage is summed up by Fricke *et al* (1986, p.489) who argued that *"few institutions are as pivotal as marriage in their implications for a broad range of social processes."* For individuals, Smith (1983, p.473) wrote that *"marriage marks maturity and permits a range of adult behaviour, including childbearing; it is a crucial rite of passage"*, while for society he stated that *"marriage creates new family nuclei and realigns households as units of consumption, savings, labour use, and production. Marriage both creates and satisfies intergenerational responsibilities and alliances."* One demographic component of all marriages is the age difference between spouses. Regardless of the cultural setting or the instigators of union formation, it has been observed that, on average, women tend to be younger than their husbands at first marriage (Dixon, 1971; Goldman *et al*, 1984; Bozon, 1991; Ní Bhrolchâin, 1992, 1997.) Despite the cross-cultural pattern in the age difference, there is much between- and within-society variation in the average age difference (Casterline *et al*, 1986; United Nations, 1990; Ní Bhrolchâin, 1992, 1997.)

1.1.1 The age difference as a hypothesised indicator

Because of its variation, the age difference has been hypothesised as associated with a variety of outcomes at both aggregate and individual levels of analyses. For example, Walsh (1972, p.193), writing about post-war Ireland, suggested that the magnitude of the age difference may have *"a major impact on the quality of Irish married life"* because of its implications for the age at childbearing and the duration of widowhood, reflecting the age at which men and women marry. Using the age difference as a hypothesised indicator of 'quality of life' may also be appropriate in the sub-Saharan African context. Because husbands tend to be much older than their wives on average in this setting (e.g. Casterline *et al*, 1986), women are more likely to become widows than their husbands become widowers (e.g. Cain, 1993). However, in some sub-Saharan African societies, women do not have inheritance rights upon the death of their husband (e.g. Cain, 1984, 1993). Associations between the age difference and marital dissolution have also been reported in the context of developed countries (e.g. Berardo *et al*, 1993). For example, in the United States it was observed that divorce and separation were more common for white couples with a large age difference in either direction; i.e. among couples where the husband was older than his wife, as well as among couples where the wife was older than her husband. However, among women who married at a relatively young age, marital instability was argued as less likely with increasingly large age differences, unless the age difference was large enough to be 'socially significant' (Bumpass and Sweet, 1972, cited by Presser, 1975).

The age difference has also been considered to be associated with a number of health outcomes. For example, it was observed from the 1976 British General Household Survey that morbidity was reported less often among married men and women where the spouses were the same age or the husband slightly older than his wife, relative to other spousal age combinations (Fox *et al*, 1979). While in the context of Sri Lanka, De Silva (1998) reported a statistically significant association between the age difference and maternal health net of a number of social-demographic factors, including the mother's age (although $p<.10$ but not $p<.05$). Specifically, greater morbidity was observed among mothers who were older than their spouses, relative to mothers who were the same age or younger than their spouses. In a study of a number of less developed countries, the age difference between sexual partners was considered as *"an important factor in explaining variations in levels of HIV infection"* (Caraël, 1995, p.97). It was suggested that a large positive age difference was associated with an increased the risk of HIV infection for women by their male partners because of men's greater opportunity for premarital sexual activity due to the later average age at marriage for men relative to women. Indeed, another study of serological concordance of wives of HIV positive men found no statistically significant association with the age difference after controlling for a number of background factors including the wife's age at marriage (Aaby *et al*, 1996).

1.1.2 The association between the age difference, age at marriage and marital status

These examples illustrate the importance of considering the association between the age difference and age at the start of the union. While often ignored, it is well established in the demographic literature that the age difference between spouses is directly related to the age at marriage (e.g. Dixon, 1971; Nawar, 1988; Bozon, 1991; Ní Bhrolcháin, 1992, 1997). For example, in the monograph 'Principles of Demography' Bogue (1969, p.329) observed that on an international basis *"the older the median age of bride at marriage, the smaller the* *discrepancy* [between the bride's and groom's age] *tends to be"*. However, within countries there is a greater association between the age difference and the age at marriage of men. For example, Goldman *et al* (1984) presented data from the United States that showed asymmetry in the age difference for males and females aged 20 to 60, and specifically that the average age difference increased with male age at marriage but remained approximately constant for females across this age range. However, it is worth noting that Goldman *et al*'s (1984) result regarding the female age at marriage is an anomaly relative to the findings of other studies, which report an initial decline in the age difference before it widens with increasing female age at marriage (e.g. Oppenheimer, 1988; Ní Bhrolchâin, 1992). The marital status of the individuals at the beginning of the union is another source of variation in the age difference (e.g. Hollingshead, 1951; Wheeler and Gunter, 1987; Ní Bhrolchâin, 1992). For example, Presser (1975) used US census data from 1960 and 1970 and estimated that the wife is at least two years older than her husband in approximately 10% of first marriages, i.e. where neither partner has previously been married. This compares to 30% of marriages where the wife has previously been married but her husband is a bachelor.

1.2 Hypotheses proposed to explain the variations in the spousal age difference

Despite the frequency with which the age difference has been considered as associated with a range of outcomes, relatively few studies have actually sought to explain the variations in the age difference at either the aggregate or individual level. Presser (1975, p.190) wrote, *"it is surprising that so little attention has been paid to such a basic demographic phenomenon as age differences between spouses and their social implications."* Some authors attribute this dearth of study to the comparatively limited availability of the male age component of the age difference, especially in the context of cross-national comparisons (e.g. Smith, 1983; Xenos and Gultiano, 1992; Cain, 1993).

Several hypotheses have been proposed to explain the variations in the age difference between and within countries. These hypotheses have often been considered in terms of the economic ideas of supply and demand (e.g. Cox, 1970b; Walsh, 1972; Boulier and Rosenzweig, 1984; Goldman *et al*, 1984; Oppenheimer, 1988; Foster and Khan, 1994). For example, one hypothesis suggests that variations in the average age difference and its distribution reflect the relative availability or 'supply' of eligible males and females (e.g.

Akers, 1967; Cox, 1970b; Dixon, 1971). Another hypothesis suggests that, at the individual level, these variations reflect preferences or 'demands' for particular age differences so that some age differences are favoured relatively more than others. A third hypothesis suggests that variations in the age difference and its distribution are a by-product of factors associated with the timing of marriage, or as Ní Bhrolcháin (1997, p.15) describes, an *"epiphenomenon of other more fundamental factors"*. This chapter now discusses these hypotheses in more detail.

1.3 The availability hypothesis

In its simplest form, the availability hypothesis suggests that variations in the average age difference and its distribution reflect the relative availability of men and women of marriageable ages. This hypothesis reflects the tendency for men to marry later than women, on average, which tends to result in a surplus of males relative to females of marriageable ages, *ceteris paribus* (Dixon, 1971).

1.3.1 The concept of a 'marriage squeeze'

The situation where there is an imbalance in the relative availability of men and women eligible for marriage to the extent that demographic consequences are hypothesised has often been referred to as a 'marriage squeeze' (e.g. Akers, 1967; Salaff, 1976; Jones, 1980; Schoen, 1983; Oppenheimer, 1988; Veevers, 1988; Greene and Rao, 1995; Timaeus and Reynar, 1998). For example, Presser (1975) argued that a marriage squeeze occurred in the United States due to the post-war baby boom, which resulted in relatively more males than females available to marry in the 1960s. Imbalances in the age-sex distribution of those able to marry may also result from sex-specific mortality. For example, there is a relative abundance of women of marriageable ages following major wars because men of marriageable ages are typically the greatest casualty (e.g. Dorjahn, 1959; McFarland, 1970). Imbalances may also result from migration. Jones (1980) for example, observed a relative shortage of females of marriageable ages in the immigrant Chinese and Indian communities in Malaysia in the late 1940s because of sex-specific trends in migration. These are examples of how sex-specific fertility, mortality, migration, and nuptiality processes may aggravate natural imbalances in the age-sex distribution of those available to marry or the 'marriage market', as this population group has often been labelled (e.g. Cox, 1970a; Goldman *et al*, 1984; Boulier and Rosenzweig, 1984; Bartiaux, 1994; Foster and Khan, 1994; Buckley, 1996).

Regardless of the cause of the imbalance, it has been suggested that *"something would have to give"* to deal with the 'marriage squeeze' (Caldwell, 1963, cited by Jones, 1980, p. 282). Responses that have been proposed include a narrowing of age differences, an increase in the incidence of polygyny, an increase in unstable marriages, and/or an increase in the proportion of women who remain single (e.g. Caldwell, 1963, cited by Jones, 1980; Greene and Rao, 1995). However, it has also been suggested that such responses may be unnecessary because men and women in the marriage market seem to adapt to the age distribution of potential partners, even in relatively extreme conditions. This hypothesis was proposed by Ní Bhrolcháin (2000) who studied birth cohort data for England and Wales from 1900 through to the late 1960s. Considerable variation in the proportions of brides and grooms who married a partner of each single-year age difference was observed and importantly, these variations mirrored the variations in the magnitude of the corresponding partner cohorts. Thus, despite large fluctuations in births, even around wartime, little evidence was found of a marriage squeeze in England and Wales. This evidence suggests that Presser's (1975) hypothesised marriage squeeze in the United States following the post-war baby boom may not actually have occurred in practice, *ceterius paribus*.

1.3.2 Rejecting the idea of a 'marriage squeeze'

This discrepancy with the popular hypothesised concept of a marriage squeeze has been attributed to the misconception that age difference preferences are inflexible, i.e. that an individual would only accept a partner precisely *x* years older or *x* years younger than them (Ní Bhrolcháin, 2000). Instead, it was inferred that preferences must be flexible in terms of an individual's willingness to accept a partner as well as an individual's willingness to be accepted as a partner (Ní Bhrolcháin, 2000). Thus, it was proposed that individuals are not simply in or out of the marriage market but that they have a conditional probability of being in the marriage market. Furthermore, there are varying degrees of flexibility such that the greater the range of acceptable age differences, the greater the flexibility. Adopting this approach means that the marriage market can be considered as operating in a *"fluid way"* (Ní Bhrolcháin, 2000, p.1). Thus, this hypothesis suggests that marriage markets can and do function effectively without the hypothesised marriage squeeze, even in societies where there is an imbalance in the marriage market as a result of sex-specific demographic processes.

1.3.3 Defining eligibility for the marriage market

This section discusses the concept of being eligible for the marriage market, and thus marriage, as eligibility is associated with marriage timing and thus the age difference. In addition to the exclusion of unmarried homosexuals and confirmed bachelors, such as priests, some heterosexual individuals of marriageable ages can not be considered as eligible for marriage (e.g. Goldman *et al*, 1984). These individuals can be considered as temporarily excluded from the marriage market, or in light of Ní Bhrolchâin's (2000) evidence, as having a small conditional probability of marriage. This is because in many societies, marriage involves leaving the natal family unit and establishing a new family unit, a process that assumes some socio-economic requirements (e.g. Oppenheimer, 1988), which some individuals of marriageable ages may not be able to meet (e.g. Dixon, 1971). As Easterlin (1973, p.20, cited by Presser, 1975) noted with respect to the 1970s United States population, *"marriage patterns are particularly affected by the economic circumstances of young males, who are typically the initiators of marriage bids"*.

In traditional societies, it has been suggested that variations in the age at marriage reflect the direction of wealth transfer between husband and wife, or at least their families (e.g. Boserup, 1970). In settings where women do not play an economic role in household production then the bride's family sometimes pays a dowry or 'bridewealth' to the groom or his family to compensate for the economic cost of marrying their daughter. In contrast, where women do contribute economically towards household production then the groom or his family sometimes pays the bride's family a 'bride price' to compensate for the bride's family's economic loss (e.g. Cherlin and Chamratrithirong, 1988; Remez, 1998). Although it is worth noting that some authors have suggested that the direction of wealth transfer can change (e.g. Lindenbaum (1981), Jones (1980), and Zhang (2000) in the context of, respectively, Bangladesh, Malaysia and China).

Regardless of the cultural setting and the direction of wealth transfer, the economic ability to marry is considered as *"an effective constraint on the number of marriages that occur"* (Walsh, 1972, p.197). This is not new idea as Malthus (1798) described socio-economic constraints as a check on population growth in his classic 'Essay on the Principle of Population'. Thus, having the economic option of an early marriage may be regarded as a luxury at the individual level (Bogue, 1969). Cleveland (1991, cited by Hampshire and Randall, 2000) observed that Ghanaian men who temporarily migrated for work could afford to marry earlier than their non-migrating counter-parts. Although, it has also been observed that depending on the length of time away, migration can result in a later age at marriage because of geographical displacement, i.e. not being available to marry (Timaeus and Graham, 1989).

1.3.4 Desiring marriage as well as being eligible for marriage

The discussion so far has assumed that men and women desire to be eligible for the marriage market as soon as possible. However, it has been suggested that the desirability of marriage depends on the availability of alternatives to marrying (Dixon, 1971). For example, Winch and Greer (1964, p.90 cited by Presser, 1975) argued in the context of more developed settings that delaying marriage for men may be desirable as well as a socio-economically necessity since a man's *"occupational position may be jeopardised by early marriage unless he can complete his training with his wife's or parents' support in the early years of the marriage"*. This suggests that a woman's economic status may enable her husband to marry at a younger age (e.g. Oppenheimer, 1988). In this situation, it is possible that age differences would be small, on average, reflecting the husband's earlier age at marriage and the delay in the timing of marriage for the wife due to the same economic constraints described above for men. Alternatively, female economic activity may facilitate early marriage for women as well as men by making marriage economically feasible for both parties in some settings (Dixon, 1971). However, a study of the United States in the early 1960s suggests that men may find women with their own careers or highly educated women unattractive, as men may be concerned that this may *"make domestic roles seem dull and unrewarding"* (Winch and Greer, 1964, p. 90, cited by Presser, 1975), although Presser (1975, p.193) added that *"support for this position may be changing."* In less developed settings, the demand for educated women in the marriage market has been considered as ambiguous because education may be considered as enhancing a woman's value but too much may be considered as negatively affecting her marriage prospects (e.g. Fricke *et al*, 1986). In contrast, at the aggregate level, women's education is frequently cited as having a positive association with the average female age at marriage (e.g. Presser, 1975; United Nations, 1987a; 1987b; Martin, 1995).

1.3.5 The marriage market as a pool of trading partners

Becker (1973; 1974) considered the marriage market as a pool of trading partners, and argued that individuals look for what they themselves lack in potential partners in an attempt to better themselves through marriage. Thus, in addition to the Malthusian theory of not marrying until one can economically afford to do so, there is also the issue of considering a potential spouse's longer-term prospects (Bogue, 1969; Schildkrout, 1983; Oppenheimer, 1988).

Increasing one's desirability may also delay the timing of marriage for an individual. Nawar (1988) suggested that this was the case in Egypt where men of a higher socio-economic status were found to marry significantly later, on average, which was attributed to the time involved in obtaining a desirable position or qualifying for a profession that requires a longer period of time out of the marriage market (see also Winch and Greer, 1964, cited by Presser, 1975). Becoming not only economically eligible for marriage but also economically desirable as a potential partner leads to the notion of inequalities in the marriage market so that some individuals are considered relatively more desirable than others (Oppenheimer *et al*, 1997). Of course, the strategy of delaying marriage in order to become more desirable in the marriage market does not guarantee finding a desirable partner (e.g. Oppenheimer, 1988). Indeed, it has been argued that a balance must be struck between delaying marriage to improve one's desirability and leaving marriage too late so as not to find a desirable partner (e.g. Boulier and Rosenzweig, 1984).

In societies where a woman's socio-economic status is defined according to her husband's socio-economic status, it has been suggested that this may motivate women to be eligible for the marriage market from an early age and to seek older men with "*visible accomplishments*" (Blake, 1974, cited by Presser, 1975, p.202). In turn, this may result in larger age differences, on average. This can be considered as an extension of Becker's (1973, 1974) hypothesis, in that where women's status is low, larger age differences are more likely, an idea that will be discussed later. Although, it has also been suggested that women may seek older partners in societies where they have the opportunity to attain their own economic status because it enables them to *"attain economic and social security without years of struggle and because they can achieve status that they can rarely acquire on their own"* (Seidenberg, 1972, cited by Presser, 1975, p.202). This leads on to the hypothesis that the age difference may in fact be a matter of female choice (Ní Bhrolcháin, 1997), an idea that is also discussed later.

1.3.6 The role of characteristics other than socio-economic status

While socio-economic status at the individual level and socio-economic conditions at the aggregate level are important in determining eligibility and desirability for marriage, other factors are also important (Schoen, 1983; Goldman *et* al, 1984; Bozon, 1991). These other factors, for example personality characteristics (Goldman *et al*, 1984), are difficult to identify because they are numerous and because it is difficult, if not impossible, to objectively conceptualise many of them at the individual level. It is often impossible therefore to identify men and women eligible for the marriage market at both the individual and aggregate levels (Smith, 1983). Furthermore, if individuals have a conditional probability of being in the marriage market, as Ní Bhrolcháin (2000) proposed, then they are not simply in or out of the marriage market. As described above, this probability depends on an individual's willingness to accept a partner of age *x* as well as their willingness to be accepted for marriage at age *x*. In addition, this idea of willingness to accept and to be accepted may apply to characteristics other than age.

1.3.7 Quantifying the marriage market

Given the difficulties associated with defining the marriage market it is not surprising that the marriage market has often been crudely quantified at the aggregate level as the ratio of men to women of a marriageable age (e.g. Akers, 1967; Dixon, 1971). For example, in a cross-national study of 28 less developed countries Casterline *et al* (1986) used the ratio of men aged 20-24 to women aged 15-19 because these age groups encompassed the median age at marriage for, respectively, men and women, in the majority of countries studied. While they acknowledged that this measure does not accurately capture marriage-market opportunities, they considered the marriage market hypothesis of *"less importance"* in explaining cross-national variations based on an analysis involving this crude measure (Casterline *et al*, 1986, p.374). More sophisticated measures of the marriage market have been developed, which in addition to considering the relative numbers of men and women, also consider suitability not just in terms of age, but also race and education, as well as the extent of competition for a partner (Goldman *et al*, 1984; Lampard, 1993). However, all of these measures assume that it is possible to define the geographical boundaries of the marriage market. For example, Goldman *et al* (1984) calculated availability ratios for the whole United States population but acknowledged that the marriage market was unlikely to operate at this level. Although, assuming that the marriage market is contained within a smaller unit area such as a metropolitan area (e.g. Goldman *et al*, 1984), or a county in the context of England and Wales (e.g. Lampard, 1993), may also be an over-simplification, especially for highly mobile populations (McFarland, 1970).

1.4 The by-product hypothesis

At the individual level, the availability hypothesis suggests that partner choice, the timing of marriage, and thus the age difference reflect the age-sex distribution of eligible men and women. However, other characteristics may be important. Cox (1970b, p.115) argued, *"this* [the age distribution of eligible persons] *is not paramount: other factors* [are] *also at work"*, because as Oppenheimer (1988, p.572) commented, *"people do not wish to marry just anyone – they want to mate assortatively"*. Thus, the by-product hypothesis suggests that variations in the distributions of the age difference may reflect preferences for particular characteristics other than age. Thus, the difference between the observed distribution of age differences and the expected distribution given the absence of a relationship between husband's age at marriage and wife's age at marriage can be hypothesised as a by-product of other forms of assortative mating.

1.4.1 Assortative and disassortative mating

The by-product hypothesis may apply as a result of homogamy, that is, positive assortative mating. This occurs when individuals with similar characteristics form partnerships, i.e. like attracts like. Conversely, the by-product hypothesis may also apply as a result of heterogamy or disassortative mating, i.e. when opposites attract (Hollingshead, 1950; Martin, 1995; Sheela and Audinarayana, 1997). Studies investigating spouse choice date back to the beginning of the twentieth century (Harris, 1912 cited by Hollingshead, 1950), and on average, tend to report more homogamy than heterogamy. For example, Hollingshead (1950) reported evidence of homogamy with respect to race and age in a study of over one thousand couples in the United States in the late 1940s. Bogue (1969) also reported this pattern in a cross-national study although he noted regional differences in terms of the amount of inter-racial marriage. For example, he reported that in Latin America, marriages occurred with *"little reference to race or ethnic origin"* (Bogue, 1969, p.644) but that in South Africa there was a high degree of positive assortative mating on these characteristics. These cross-national variations in the amount of inter-racial marriage may reflect the availability of marriage partners, the relative size of the racial groups, and the level of racial prejudice and residential segregation (Schmitt, 1965, cited by Bogue 1969).

1.4.2 The opportunity to marry assortatively

A more consistent pattern of homogamy has been observed cross-nationally with respect to the educational backgrounds of spouses (e.g. Epstein and Guttman, 1983; Martin, 1995; Kalmijn, 1998). Assortative mating with respect to education clearly has implications for age at marriage. For example, if an individual desires a partner with a university education then this implies that their partner will have a certain minimum age, assuming completion of their university education before marriage. In addition, Warren (1966, cited by Epstein and Guttman, 1983, p.250) argued that *"education sorts people into demographically separate groups at marrying age."* This refers to the practice of endogamy, whereby individuals choose their marriage partners from within their social group (e.g. McFarland, 1970; Bozon and Heran, 1989; Kalmijn, 1998) or 'social circle' (Henry, 1972). This suggests that, on average, individuals do not choose someone with a university education over someone without a university education, rather they are socialised or sorted into groups. Thus, if an individual has a university education then it is more likely that their partner will also have a university education. More generally, this hypothesis suggests that an individual may prefer a partner with a particular characteristic but if they do not have the opportunity to encounter individuals with this characteristic then the probability of having such a partnership is low. The hypothesis that individuals in the marriage market have a conditional probability of selecting individual x as a partner and being selected for marriage by partner x is therefore also conditional on meeting such an individual (Ní Bhrolcháin, 2000).

1.5 The preferences hypothesis

In the same way that individuals may have preferences regarding a partner's characteristics such as their education and religion, there is a hypothesis that proposes that average age differences and distributions of age differences reflect preferences for a partner of a particular age or rather, as Ní Bhrolcháin (2000) proposed, age *range*. It has been suggested that individuals' preferences may reflect pressure from society to conform to norms about what is and what is not an acceptable age for a partner. For example, Presser (1975, p.196) argued that *"just as there are norms about the age at which women and men should marry, there are norms about age differences between spouses"* (see also Hollingshead, 1951; Jones, 1980). Furthermore, Richman (1977, cited by Atkinson and Glass, 1985) argued that the more traditional a society, the more important it is for individuals to adhere to norms regarding the age difference. Akers (1967, p. 908) assumed the existence of such norms when he suggested that women who delayed marriage may ultimately have had to marry a man of a *"less appropriate age"* else remain unmarried. Park and Cho (1995, p.74) also used this terminology with respect to imbalances in the Korean marriage market in the 1980s and 1990s when they claimed that some men were unable to find wives of an *"appropriate age"*.

1.5.1 The significance of age

One starting point in determining what constitutes an appropriate age for a spouse, and thus age difference, is to consider the symbolism and significance of age. Neugarten *et al* (1965, p. 710) described that *"in all societies, age is one of the bases for the ascription of status...age norms and age expectations operate as prods and brakes upon behaviour, in some instances hastening an event, in others delaying it."* However, in some societies, typically more traditional societies, little significance is attached to chronological age because rites of passage are considered more influential (Parsons, 1940; 1942; Dorjahn, 1959; Brass *et al*, 1968). Although, Bozon's (1991, p.121) description of the significance of age suggests that it is a subjective concept in all societies: *"age is no simple, unchanging and objective reality, but a complex socio-historical construction that institutes classifications, comparisons, oppositions."* For example, Presser (1975, p.190) argued that in the United States in the 1970s age differences where the wife was older than her husband were considered *"deviant"* and *"far more problematic"* than age differences where the husband was older. However, Feng and Quanhe (1996) estimated that one-quarter of ever-married women had a negative age difference from China's 1988 'Two-per-thousand Fertility and Birth Control Survey', which suggests that such age gaps are not 'deviant' in this society. Of course, regardless of whether or not significance is attached to chronological age, a couple's age difference may not be discernible to those other than the individuals themselves unless its magnitude is such that it is obvious that the husband is older than his wife or *vice versa* (Ní Bhrolcháin, 1997). Thus, individuals may have preferences for relatively extreme age differences in order to make a statement for either or both parties in the union (Dorjahn, 1959). Despite the popularity of the assumption that individual preferences are influenced by societal norms, it has been argued that, in actual fact, there is no evidence either for or against the influence of norms in determining preferences so that *"it seems a rather tenuous basis on which to make a strong assumption central to the functioning of marriage markets"* (Ní Bhrolcháin, 2000, p.27).

1.5.2 Examining the preferences hypothesis

Very few studies have examined the preferences hypothesis. From a literature review on mate selection, Jensen (1978, cited by Epstein and Guttman, 1983) reported that age was the characteristic which shows the greatest degree of similarity between spouses. However, it is important not to assume causation from this association and interpret this as evidence for the preferences hypothesis as these observations may simply reflect the age-sex structure of those in the marriage market. For example, from a study of age difference preferences in France, Bozon (1991) observed that negative age differences were not uncommon, especially among women who married relatively late, reflecting the age-sex structure of those available in the marriage market when these women were marrying. However, Bozon (1991) also reported that these women actually preferred a positive age difference. Thus, conclusions regarding age difference preferences based simply on the observed distribution of age differences would have been false.

It is likely that one of the main reasons that the preferences hypothesis has received relatively little examination is because of a shortage of explicit data on age preferences. One study that used such data considered the age range of partners whom British male and female dating agency clients were willing to accept (Ní Bhrolcháin, 1997). It was observed that the age range reported by women was much narrower than that reported by men, and that while both sexes reported that they preferred a positive age difference, *"men appear more willing to forego a positive age difference than women do to accept a negative one"* (Ní Bhrolcháin, 1997, p.11). While this sample may not be representative of the British population, similar patterns were reported from a study of a representative, general population sample of couples in France (Bozon, 1991). These results suggests that women are relatively more choosy than men regarding age difference preferences, and thus that variations in the age difference may reflect female rather than male choice (Ní Bhrolcháin, 1997). This is not a new hypothesis. For example, Walsh (1972, p.202) suggested that in Ireland from the mid-1950s to the late 1960s women were reluctant to marry men too far removed from them in age in either direction, and concluded that women's preferences have been *"a dominant force in accounting for the declining inequality* [in age at marriage]." This hypothesis has also been implied in the context of a less developed society. In a qualitative study descriptively titled *"Sugar daddies and gold diggers"*, it was reported that in Accra, Ghana, young, white-collar working women preferred and chose much older, affluent men because these men could indulge the women with money and material gifts while the girls could provide the men with *"socio-sexual services"* (Dinan, 1983, p.356).

1.5.3 The women's status hypothesis

A more common hypothesis, which is related to the preferences hypothesis, suggests that the age difference is negatively associated with the status of women at both the aggregate- and individual levels (e.g. Presser, 1975; Mason, 1985, 1995; Atkinson and Glass, 1985; Casterline *et al*, 1986; Sokona and Casterline, 1988; Nawar, 1988; Lesthaeghe *et al*, 1989; Jejeebhoy, 1991; Cain, 1993; Balk, 1994, 1997; Greene and Rao, 1995; Feng and Quanhe, 1996; Niraula and Morgan, 1996; Villarreal, 1998.) Despite its popularity, many authors have simply assumed the 'women's status hypothesis' without examining its validity. For example, at the individual level, Barbieri and Hertrich (1999) assumed that the age difference reflects the extent of equality between spouses in their study of fertility in the context of seven sub-Saharan African countries.

In contrast, Casterline *et al* (1986) examined the women's status hypothesis at the aggregate level by examining the association between the average age difference and girls' enrolment in primary schools in 1960, which they considered as a frequently-used indicator of women's status. The study concluded that *"variations in the age difference distributions between societies can be directly interpreted in terms of two sets of factors: kinship structure and the status of women, themselves closely linked"* (Casterline *et al*, 1986, p.374). However, as discussed above, Casterline *et al* (1986) crudely controlled for the relative availability of eligible men and women with the sex ratio. Thus, the conclusion regarding the women's status hypothesis at the aggregate level may have been incorrectly deduced.

More recently, a cross-national study by Cain (1993) observed a positive, albeit weak association between the age difference, which was hypothesised as an indicator of patriarchal structure, and in turn, the total fertility rate and an indicator of son-preference. However, like the majority of studies of the age difference and its hypothesised association with women's status (including Casterline *et al*, 1986), Cain's (1993) analysis did not control for age at marriage. This is despite the established association between the age difference and age at the start of the union (e.g. Bogue, 1969; Dixon, 1971; Nawar, 1988; Bozon, 1991; Ní Bhrolcháin, 1992, 1997), and that the female age at marriage is often considered as an indicator of women's status (e.g. Mason, 1985; Chowdhury and Trovato, 1994; Balk, 1994). Indeed, one of the few studies that did examine the association between the age difference and indicators of women's status net of the wife's age at marriage observed no statistically significant evidence in support of the women's status hypothesis at the individual level in the context of Turkey (Remez, 1998). Thus, relative to the numerous studies that have simply assumed the women's status hypothesis, only a few studies have examined this hypothesis, and of these, even fewer have considered the three-way association between the age difference, age at marriage and the status of women.

1.6 Conclusions

This chapter has attempted to review the main hypotheses that have been proposed in the literature to explain the variations in the age difference both cross-nationally and within countries. The three main hypotheses discussed were referred to as the availability hypothesis, the by-product hypothesis, and the preferences hypothesis. Although the hypotheses were discussed in turn, in practice the hypotheses are not mutually exclusive as has been evident from some of the examples quoted. While the availability hypothesis and the by-product hypothesis complement one another, it has been suggested that the availability hypothesis and the preferences hypothesis may *"co-exist either with or without tension, depending on the relationship between age difference preferences and marriage market opportunities"* (Casterline *et al*, 1986, p.354). Although, this idea of tension contradicts Ní Bhrolcháin's (2000) idea of flexibility in the marriage market as discussed above. Instead, the availability hypothesis and the preferences hypothesis may also need to be considered as complementing one another. Despite the frequency with which these hypotheses have been assumed in the literature, relatively few studies have examined these hypotheses. Indeed, few have identified the extent of variation in the age difference, at either the aggregate or individual levels of analysis.

1.7 Thesis objectives

This thesis has two main objectives. The first objective in Chapters 3 and 4 is to describe the extent of variation in the average age difference between countries and in the distribution of age differences within countries, where data are available. Chapter 3 begins by using aggregate-level data from the third United Nations compiled Women's Indicators and Statistics database for 93 less- and more-developed countries around the world to try and paint a global picture, albeit limited by data availability. Given the greater variation in the average age difference between and within less developed regions in comparison to the relatively more developed regions (Casterline *et al*, 1986; United Nations, 1990), Chapter 4 then uses individual-level data for as many countries as possible as collected by the Demographic and Health Surveys from five, less developed regions (Central and South America, West Asia and North Africa, South Asia, Central, East and South Africa, and West Africa).

In addition to examining variations in the age difference, the often-overlooked association between the age difference and age at marriage will be examined, as this association may be important in understanding the variations in the average age difference and its distribution. This thesis does not examine the association between the age difference and the marital status of individuals at the start of the union since, in most cases, the age difference is calculated from data that refer to the age at first marriage. This reflects the dearth of data on marriages of a higher order on an international basis. This in turn reflects the significance of first marriage especially for women in terms of exposure to the risk of childbearing, which still exists in many societies today (e.g. Cox, 1970a; Bongaarts, 1978; McCarthy, 1982). This is in addition to a general scarcity of male nuptiality data (e.g. Smith, 1983; Xenos and Gultiano, 1992; Cain, 1993).

The second objective is to empirically examine the hypotheses discussed in this chapter, starting at an aggregate-level of analysis and proceeding to an individual-level of analysis where appropriate, and again, where the data permit this. Chapter 3 begins by examining the women's status hypothesis at the aggregate level, specifically that the average age difference reflects the status of women. Chapter 4 examines the availability hypothesis, that variations in the average age difference and the distributions of the age difference reflect the age-sex structure of those marrying. Chapter 5 examines the by-product hypothesis using individual-level data for each of the 30 less developed countries. Chapter 6 concludes the analysis by examining the women's status hypothesis at the individual level. This hypothesis is examined using individual-level data from the 1995 Egyptian Demographic and Health Survey, which unlike the other DHS surveys, included an additional module on the status of women.

In addition, there is a focus on the extent of age misreporting given the extensive literature on this aspect of data quality (e.g. Carrier, 1959; Brass *et al*, 1968; Nagi *et al*, 1973; Ewbank, 1981; Morah, 1985). In order not to distract from the main emphasis of this thesis, the focus on age misreporting is reported in Chapter 2, which follows, and which deals with data issues and methods. Chapters 3 through 6 present the results of the examinations of the hypotheses. Chapter 7 concludes by drawing together these results.

Chapter 2:
Data issues and methods

2.1 Overview

This chapter discusses the data available and the methods employed in this thesis to identify the variations in the age difference and to examine the hypotheses outlined in Chapter 1. To begin, the availability of data at the aggregate-level is discussed, and then the thesis's aggregate-level data source, the Women's Indicators and Statistics database, is introduced. The first hypothesis to be examined in Chapter 3 is that the age difference reflects the status of women at the aggregate level so there is a discussion of some of the conceptual and measurement issues surrounding women's status. The interpretation of women's status and methods of analysis used for this aggregate-level analysis are then described. This chapter then introduces the thesis' individual-level data source, the Demographic and Health Surveys, which are used to identify variations in the age difference between and within 30 less developed countries in Chapter 4. The quality of age data from these surveys is examined before describing the methods of analysis used in Chapters 4 and 5 to examine the marriage market and by-product hypotheses. Finally, this chapter discusses the data issues and methods of analysis of the individual-level data from the 1995 Egyptian Demographic and Health Survey presented in Chapter 6, the purpose of which is to examine the hypothesis that the age difference reflects the status of women at the individual-level.

2.2 Cross-national variations in the age difference using aggregate-level data

As discussed in Chapter 1, a number of studies have assumed that the age difference reflects the status of women at both the aggregate and individual levels. This study utilises aggregate level data from the Women's Indicators and Statistics Database to investigate this hypothesis at the aggregate level for 93 less developed and more developed countries. This data source and the methods employed are discussed now.

2.2.1 The Women's Indicators and Statistics database

At the time of writing, there are very few aggregate-level demographic data sources that provide data for countries in all global regions. For example, the Demographic and Health Surveys only collect data from less developed countries, while the Statistical Abstract of Latin America is a data compendium specific to countries in Latin America. As far as international data on the age at marriage is concerned, the United Nations (1990) 'Patterns of Marriage: Timing and Prevalence' has been frequently cited (e.g. Xenos and Gultiano, 1992; Meekers, 1992; Isiugo-Abanihe, 1994; Parrado and Tienda, 1997; Ní Bhrolcháin, 1997). The UN's (1990) estimates are considered reliable since they are based on census data or other large-scale representative surveys although, some of its most recent estimates are as much as 40 years old.

In 1997, the UN published the third edition of the Women's Indicators and Statistics database (hereon referred to as WISTAT3). This is a compilation of national estimates available for 212 less developed and more developed countries as at July 1994. The data cover a wide range of topics including population composition, marriage, health, education, economic activity, and political involvement, and, where possible, these data are presented separately for males and females. One such variable is the average age at marriage, which is estimated via the singulate mean age at marriage (hereon referred to as the SMAM).

2.2.2 The singulate mean age at marriage

The SMAM was devised by Hajnal (1953) and is estimated on a synthetic cohort, based on the proportion of the population unmarried at any one time. It is a standardised measure and, since it is unaffected by the age structure of the population for which it is calculated, it is useful for cross-national analyses. Since it uses cross-sectional data such as censuses and surveys, another advantage of this measure is that the average age at marriage can still be estimated for countries where vital registration systems are unreliable (e.g. Agarwala, 1957).

The principle of the SMAM is to estimate the total years lived by single persons up to the age of 50 by those who marry by age 50, then to divide this by the number who marry by age 50. This obtains the average number of years single among those who marry by age 50, which is equal to the average age at first marriage since first marriage marks the end of being single. The SMAM method makes two assumptions (1) that individuals who marry do so before exact age 50, and (2) that marriage rates have been approximately constant in the past. It is important to question the extent to which the second assumption is valid since the age at marriage of both males and females has been fluctuating, with an increase in female age at marriage in many less developed countries (e.g. UN, 1990; Isiugo-Abanihe, 1994; Ní Bhrolcháin, 1992; 1997). Thus, it is possible that the SMAM may under-estimate the true average age at marriage. Furthermore, since this analysis estimates the aggregate age difference as the male SMAM minus the female SMAM, if the changes in marriage timing have been different for men and women then this may also distort the estimate of the age difference (Singh, 1992). Despite these issues, the SMAM is the only estimate of marriage timing that is available on a sufficiently wide basis to permit a cross-national study of the spousal age difference. Further details regarding conceptual and methodological issues of using the SMAM are given in Pollard *et al* (1990) and Shryock and Siegel (1976).

2.2.3 Data availability and selection from the WISTAT3

WISTAT3 presents data for 212 countries, territories and colonies. However, for some of these countries, especially the smaller territories and colonies, and some less developed countries, data are not available for some indicators. Unfortunately for this study, this includes the SMAM. Furthermore, since some of the estimates of age at marriage are obtained from the UN's (1990) compendium mentioned above, some estimates are quite out of date. In order to try to present a more up-to-date picture of cross-national variations in the average age difference, this thesis imposes some selection criteria for the inclusion of countries in this analysis. A total of 46 countries with small populations, defined here as less than 0.5 million people, are excluded because of limited data availability. A further 65 countries are excluded because either the SMAM is not available for the same year for both sexes, or the data were collected before 1980. Table 2.1 shows how the WISTAT3 sample is reduced to a sample of 93 countries for analysis in this thesis, by UN region.

Table 2.1 WISTAT3 sample size by United Nations region

	United Nations region[1]				
	Africa	**Asia & Pacific**	**Latin America & Caribbean**	**Developed countries**	**All countries**
	N (%)	**N (%)**	**N (%)**	**N (%)**	**N (%)**
Total number of countries in WISTAT3[2]	55 (25.9)	65 (30.7)	39 (18.4)	53 (25.0)	212 (100.0)
Exclusion criteria[3]:					
Country's population size in WISTAT3:					
Not stated[4]	1 (1.8)	2 (3.1)	0 (0.0)	5 (9.4)	8 (3.8)
< 0.5 million[4]	6 (10.9)	19 (29.2)	15 (38.5)	6 (11.3)	46 (21.7)
Age at marriage data:					
Not available for males and/or females[4]	7 (12.7)	15 (23.2)	0 (0.0)	13 (24.5)	35 (16.5)
Pre-1980[4]	17 (30.9)	4 (6.2)	5 (12.8)	1 (1.9)	27 (12.7)
Year to which the age at marriage data refer is different for men and women[4]	3 (5.5)	0 (0.0)	0 (0.0)	0 (0.0)	3 (1.4)
Number of countries in this study's sample[4]	21 (38.2)	25 (38.5)	19 (48.7)	28 (52.8)	93 (43.9)
Regional percentage distribution[5]	21 (22.6)	25 (26.9)	19 (20.4)	28 (30.1)	93 (100.0)

Notes for Table 2.1

1. Regions are defined according to the UN Statistical Division. See Appendix 2A for country listings by UN region
2. Numbers in parentheses refer to the per cent of all 212 countries in WISTAT3
3. Some countries are excluded for more than one reason. The numbers in Table 2.1 refer to the number of countries remaining after removing countries according to each exclusion criteria in turn.
4. Numbers in parentheses refer to the per cent of all countries in WISTAT3 for the region
5. Numbers in parentheses refer to the per cent of all 93 countries in this study

The last two rows of Table 2.1 show that this study's sample of 93 countries is not entirely representative of the regional distribution of the 212 countries in the WISTAT3 dataset, according to UN's definition of region. Developed countries and countries in Latin American and the Caribbean are slightly over-represented, while countries in the African and Asian regions are slightly under-represented in this sample. Having identified countries with 'valid' age at marriage data for this analysis, it is necessary to apply the same criteria to the indicators of women's status. First, it is important to address some of the conceptual and measurement issues surrounding women's status.

2.2.4 Conceptual and measurement issues surrounding women's status

Over the past few decades, numerous studies have involved discussions about the status of women. Despite the topic's popularity, there remain many debates as to the specific meaning of the concept of women's status. According to the frequently-cited discussion on the subject by Mason (1985, cited by, amongst others, Smith, 1989; Malhotra, 1991; Vlassoff, 1992; Simmons and Young, 1995; Niraula and Morgan, 1996; Govindasamy and Malhotra, 1996; Schuler *et al*, 1997; Parrado and Tienda, 1997; Riley, 1997), three of the most common interpretations of women's status consider it in terms of (1) women's position, (2) women's roles, and (3) women's autonomy.

Women's status seems to be most frequently interpreted and discussed in terms of women's position or social standing (e.g. Caldwell and Caldwell, 1993; Balk, 1994; Mhloyi, 1994; Mason, 1995). This interpretation relates the position of women to the position of men (e.g. UN, 1987a; Adepoju and Oppong, 1994; Riley, 1997), or, in some instances, the position of other women. For example, the UN (1984) compares women of different ages, and also of different social groups in this respect. However, there is no consensus as to which age, social group, or population should form the reference. For example, Schildkrout (1983) reports how, in some Islamic settings, despite strict cultural taboos against women participating in work beyond the domestic setting, there nonetheless exists an elaborate network of exchange and income generation among women and children. In comparison to the position of men, women may be considered as having low status in this setting, but comparisons between women would identify much variation in terms of their position in this 'hidden trade'. Such comparisons are not only culturally-specific but also time-specific. For example, in Britain, the position of women today has dramatically changed by comparison with the position of women in the late nineteenth century or even as recently as a couple of decades ago. Perhaps this is one reason why the position of women is frequently compared to the position of men since such a comparison can be applied to the same society and to the same point in time. Although, since this interpretation measures inequality, this may have implications for obtaining accurate data since as van de Walle and van de Walle (1993, p.65) note *"no society will freely admit that the status of its women is low"*.

Another interpretation of women's status is in terms of women's roles or social expectations (e.g. Mason, 1987; Oppong, 1995; Johnson and Turnbull, 1995; Riley, 1997). This interpretation suggests that women's roles are socially ascribed in that there are expected behaviours with associated status. For example, it has been argued that marriage and childbearing are central events for most Egyptian women (Nawar, 1988; Sokona and Casterline, 1988; Mahran *et al*, 1995), since *"the family unit is central in Arab societies. Societal recognition and support systems revolve around the roles of women as wives and mothers"* (Rashad, 2000, p. 94). While this comparatively traditional emphasis may be interpreted as women having low status in relatively more developed settings, in Egypt women are likely to have considerable autonomy at least in terms of reproductive issues because these issues fall within the *"socially prescribed female domain"* (Govindasamy and Malhotra, 1996, p.329). Thus, as with the interpretation of women's position, this interpretation also faces challenges of time and cultural specificity. Furthermore, while some may argue that it escapes gender specificity, it is necessary to question the extent to which these socially ascribed roles are truly independent from gender issues in a patriarchal society.

Women's status is sometimes considered as woman's autonomy in terms of her control over her resources and her ability to make and act upon her own decisions. Dyson and Moore (1983, p.43) define autonomy as *"the ability – technical, social, and psychological – to obtain information and to use it as the basis for making decisions about one's private concerns and those of one's intimates."* Although others stress the need to distinguish between aspects of autonomy concerned with women's access to resources and aspects of autonomy concerned with women's control over resources (e.g. Mason, 1993; Simmons and Young, 1996). For example, Hagan (1983) describes the traditional family system of Effutu in the fishing community of Winneba, Ghana. On the one hand, women control the financial resources at the household and community level, as they are responsible for preparing and selling the fish that their husbands catch. However, the women must give their husbands the money from their sales which, if it is not in line with the amount that other husbands receive, is considered grounds for divorce. The husband then gives his wife money for family and domestic expenses. Thus, while women in this setting may have autonomy in terms of controlling financial resources at the community level, these women have limited access to the resources at the individual-level. Whether or not the status of these women is considered high depends therefore on how autonomy is interpreted. Indeed, others do not consider that autonomy is even related to women's status. For example, Caldwell and Caldwell (1993, p.123) argue that *"women's autonomy is not the same as women's status, at least as measured by the potential for respect or reverence; indeed, the opposite may be closer to the truth"*. Van de Walle and van de Walle (1993, p.61) also make this distinction, *"our topic is not really the status or the condition of women per se, but the extent to which women are making decisions which affect their own fertility"* (see also Niraula and Morgan, 1996).

In practice, it is often impossible to disentangle these three interpretations of women's status, and this includes categorising hypothesised indicators according to these interpretations. Yet this has not prevented researchers from trying. For example, Powers and Salvo (1982, p.22) refers to their study's indicators of the status of women as *"conceptually manageable"*, but this simply means using social, economic and demographic labels rather than actually defining women's status. Similarly, Caldwell and Caldwell (1993) considered women's status as having eight broad elements, including kinship, education, social status, work, income, and three autonomy elements. This method of categorisation was then used in an attempt to measure women's position independently of their husband's 'position' (Caldwell and Caldwell, 1993). However, the extent to which this can really be achieved is questionable. For example, surveys such as the Demographic and Health Survey sometimes do not distinguish between the husband and wife, rather questions may talk more generally in terms of the household.

Regardless of the interpretation of women's status employed, it is difficult to find one indicator that is appropriate for all settings and for which reliable, comparable data are available. Table 2.2 shows a compilation of 34 indicators that are frequently assumed to be associated with women's status in the socio-demographic literature, and the direction of this hypothesised association, as featured in Mason (1985, 1986). Included in this compilation is the spousal age difference, which, as discussed earlier is hypothesised as having a negative association with the status of women.

Table 2.2 Indicators of women's status commonly used or mentioned in the social demographic literature

Indicator of women's status	Hypothesised association with women's status
Demographic	
Female - male mortality rates	-
Female age at marriage	+
Average husband - wife age difference	-
Parents' preference for male children	-
Kinship	
Purdah (female seclusion)	- (?)
Levirate (enforced marriage of widows to husband's brother)	- (?)
Polygamy (multiple wives)	?
Conjugal family households	+ (?)
Emphasis on lineage	- (?)
Female property inheritance	+
Village exogamy of females (out-marrying)	+
Patrilocal post-marital residence	-
Dowry	-
Arranged marriages	-
Cross-cousin marriages	-
Emphasis on virginity of brides	-
Pre- or post-marital sexual double standard	-
Emphasis on women's sexuality, youthfulness	-
Male right to divorce wife without her consent	-
Egalitarianness of the husband-wife relationship	+
Male feeding priority	-
Extended-kin support for widows and divorcees	+
Economic	
Female employment 'opportunities'	+
Female labour force participation	+
Exclusion of women from extra-domestic activities	-
Concentration of women vs. men in informal economic sector	-
Occupational segregation of the sexes	-
Sex differences in wages or earnings	-
Sex differences in the amount of leisure time	-
Female education	+
Female under-employment or unemployment rates	-
Women's work 'commitment' (measured variously)	+
Women's access to credit	+
Women's access to non-familial supports	+

Source: Mason, K. O. (1985, p.26) *The status of women. A review of its relationships to fertility and mortality.* The Rockefeller Foundation: New York. Also in Mason, K. O. (1986, p. 295) The status of women: Conceptual and measurement issues in demographic studies. *Sociological Forum*, vol.1, no.2, pp. 284-300.

Clearly, some of the indicators in Table 2.2 are more useful for cross-cultural studies, such as female age at marriage, education, and labour force participation (Pinnelli, 1993; Kishor and Neitzel, 1996), relative to culturally specific indicators such as the practice of Purdah, exogamy, and male feeding priorities. Furthermore, as Table 2.2 shows, the expected association between some of the more culturally sensitive indicators and women's status is uncertain. This begs the question: how useful are attempts to measure women's status if the concept or at least some of its measures are culturally limited? This implies that some indicators are more robust than others especially in the context of cross-cultural studies. Furthermore, there are three sets of factors in Mason's (1985, 1986) table: demographic, kinship and economic. While these factors may or may not be related, these three sets of factors suggest that women's status is multidimensional, such that it can not be captured with just one indicator.

It is also important to note that many of the indicators hypothesised as indicators of women's status at the aggregate level are also considered as indicators of the level of development in a society (Dixon, 1971; Lesthaeghe *et al*, 1989; Obermeyer, 1992). For example, Dyson and Moore (1983) used the proportion of the female adult population who were literate as an Indian state-level indicator of women's status. However, the UN's Human Development Index, which ranks countries according to their level of socio-economic development is derived from four indicators, Gross Domestic Product, life expectancy, school enrolment, and adult illiteracy (UN, 1999). The association between women's status and socio-economic development has been well documented (Boserup, 1970; Lesthaeghe *et al*, 1989; Mazur and Mhloyi, 1994; Mhloyi, 1994; Simmons and Young, 1995; Ogbuagu; 1996). In an attempt to measure women's status 'net' of the level of development within a society, women's status has sometimes been defined as women's position relative to men as noted above. Mason's (1985, 1986) compilation above gives examples of this, including the arithmetic gender difference in mortality, age at marriage (i.e. the spousal age difference), wages, leisure time, and the concentration of women versus men in the informal economic sector.

2.2.5 The interpretation of women's status for the aggregate analysis

Given the continuing debate surrounding how women's status is conceptualised and how indicators are defined, Chapter 3 attempts to measure women's status in several different ways. Firstly, three indicators that refer specifically to women are selected. These are:

1. Adult illiteracy: The population aged 15 and over that is illiterate or semi-literate as a percentage of the total population aged 15 and over.
2. School enrolment: The population aged 12-17 enrolled in school as a percentage of the total population aged 12-17.
3. Life expectancy at birth: The number of years a newborn would be expected to live according to the age-specific mortality rates prevailing at the time of their birth.

Since the analysis in Chapter 3 is a cross-national comparison, the indicators must be appropriate for all settings. It is for this reason that indicators with subjective definitions are avoided, such as economic activity (Fapohunda, 1983; Adepoju and Oppong, 1994). It is also for this reason that educational indicators are often used in the literature (Mason, 1985, 1986; Kritz and Gurak, 1989; Lesthaeghe *et al*, 1989; Jejeebhoy, 1996), and why two education indicators are used in Chapter 3's analysis. While associated, one refers to the educational level of married women and the other, the contemporary society's attitude towards female education. Education can therefore be considered as reflecting both opportunity and realisation of women's status (e.g. Adepoju and Oppong, 1994; Mason, 1995).

Secondly, the arithmetic gender difference for each of the three indicators is used as a tentative proxy for the position of women relative to men in a country. Thus, controlling for the absolute level of an indicator can be considered as a tentative proxy for women's status net of socio-economic development within a society. By definition, it is necessary to have data available separately for men and women, which is another reason that the three indicators listed above were selected from WISTAT3. In order to maintain consistency throughout the analysis, the gender difference in these three indicators is calculated as the female variable minus the male variable, although the age difference is calculated as the male age minus the female for convention. To ensure valid comparisons, the same time-constraints are applied to the women's status indicators that are applied to the age at marriage data.

Thirdly, it is recognised that many indicators hypothesised as associated with women's status are correlated and so Chapter 3 also considers women's status multidimensionally, as an 'umbrella' concept. Factor analysis is employed to identify several common factors from a number of correlated indicators of women's status. Since the analysis in Chapter 3 is based on 93 countries, it is considered reasonable to use twelve indicators in the factor analysis (Child, 1970). Indicators measuring the gender difference are excluded since they are created from the indicators that refer only to females, which are already included in the factor analysis. This is because there is automatic correlation between the gender difference and the female-specific indicators, and this would artificially create factors (Child, 1970).

All but one of the twelve indicators for the factor analysis are extracted from WISTAT3. The exception is the proportion of the population of a country who are Muslim. Given the discussion in the literature about the importance of cultural context in understanding the status of women (Govindasamy and Malhotra, 1996), it is considered important to include a more socio-cultural indicator. However, since this is a cross-national and cross-cultural analysis, there are few socio-cultural indicators that apply cross-nationally (Fricke *et al*, 1986). The proportion of the population who are Muslim may be considered as one such indicator as the Muslim faith is characterised by early, and often arranged marriage, post-marital patrilocal residence, Purdah (the Muslim custom of veiling and segregating women, e.g. Shah and Bulatao, 1981; Bisilliat, 1983), and the sanction of polygamy (e.g. Obermeyer, 1992; Islam and Ahmed, 1998). It is acknowledged that there is much variation in the practice of this faith, and thus, the extent to which these characteristics apply at different levels of analysis (Chamie, 1986; Sachedina, 1990; Obermeyer, 1992). However, this is not the first study that has attempted to control for this socio-cultural characteristic (e.g. Balk, 1994; Schuler *et al*, 1997).

Estimates of the proportion of each country's population who are Muslim are obtained from 'The World Factbook' (Central Intelligence Agency, 1999). It is published annually by the United States government and provides a profile of every country of the world, including estimates of socio-demographic information such as the religious make-up of countries. These estimates are not considered without bias but, at the time of analysis, these were the only estimates available on an international basis. In addition to the absence of socio-cultural data, data are missing from WISTAT3 for some of the variables used in this aggregate analysis and so it is necessary to refer to the original data sources to supplement the WISTAT3 where possible. Appendix 2B gives the proportion of the sample with missing data and the data sources used to supplement WISTAT3.

Chapter 3 uses the principal components method of factor extraction to determine the factors. In deciding the number of factors that are necessary to adequately represent the data, the proportion of the total variance that each factor explains is examined. Several procedures are described in the literature for deciding the number of factors to use in a model (Norušis, 1992). The criterion that Chapter 3 uses states that only factors that account for eigenvalues, or variances, greater than one should be included (Child, 1970; Ehrenberg, 1975; Manly, 1986; Norušis, 1992). The initial solution of a factor analysis is usually difficult to interpret because the factor loadings all tend to be quite large, especially for the first factor since the first factor explains most of the variance, and also because variables may have large factor loadings for more than one factor. In an attempt to aid interpretation, the factors are rotated to redistribute the explained variance so that each factor has large loadings for just a few variables (although this does not affect the goodness of fit of the solution). There are two types of rotation: orthogonal, which assumes that the factors are not correlated, and oblique, which assumes that the factors are correlated. Given the literature on the association between different dimensions of women's status, it seems reasonable to assume that the factors are likely to be correlated. Thus, the factor analysis in Chapter 3 uses the oblimin oblique method of rotation to obtain factor loadings as this aids interpretation. Factor loadings greater than ±0.3 are considered significant (Child, 1970).

2.2.6 Methods of analysis used in Chapter 3

The statistical packages SPSS version 9 (Norušis, 1992) and JMP (Sall *et al*, 2001) are used for the descriptive analysis of the aggregate-level age difference data. Chapter 3 presents all estimates to one decimal place since the SMAMs are given to one decimal place in WISTAT3. To begin, boxplots are used to display the extent of variation in the age difference between and within geographical regions as used elsewhere for this purpose (e.g. Ní Bhrolcháin, 1997.) Boxplots are created in SPSS and show the median, lower and upper quartiles, outlying values (those values between 1.5 and 3 times the inter-quartile range from the lower or upper quartile), and extreme values (those values greater than three times the inter-quartile range from the lower or upper quartile), for each region. The median and the inter-quartile range are used in preference to the mean and standard deviation as normality can not be assumed for the regional distributions. These methods are then used for the analysis of the aggregate-level age at marriage data.

As identified in Chapter 1, the association between the age difference and age at marriage is seldom examined at the aggregate level so Chapter 3 will examine this association at this level of analysis. Scatterplots are used to show the association between the age difference and the average age at marriage for males and then females. It is not possible to use correlation coefficients to measure the extent of association between the age difference and the average age at marriage of females since a curvilinear association is identified. Instead, linear regression is used with an additional term, female age at marriage squared, to take account of the curvilinear association at the aggregate-level. The adjusted R-square is used to estimate the proportion of the variation in the age difference that is explained by the average age at marriage of males and females, in turn, at the aggregate-level.

Before considering the association between the age difference and the hypothesised indicators of women's status, boxplots are again used to display the extent of variation between and within geographical regions for each of the indicators. Scatterplots are then used to show the extent of association between the age difference and each indicator of women's status. JMP is used to apply a smoothing spline to each plot to highlight any association. The spline consists of a set of third degree polynomial segments that are linked together so that the resulting curve is continuous and smooth at the knots. As the value for the smoothness parameter, lambda, tends towards zero, the error term has more weight and the fit of the spline becomes more curved. Conversely, as the value of lambda increases, the curve tends towards a straight line, but the amount of variation that can be explained decreases. Thus, splines can account for more variation than fitting a straight line or for example, quadratic curve. The choice of lambda is subjective. Using lambda equal to 100 is necessary for most of the analyses of the unidimensional indicators, but lambda equal to ten is sufficient for the analyses of the multidimensional indicators. Further details regarding splines are given in Sall *et al* (2001).

The average age at marriage of females is often considered as an aggregate-level indicator of the status of women (e.g. Smith, 1983; Mason, 1985; Singh and Samara, 1996). Thus, Chapter 3 examines the associations between the average age difference and the hypothesised indicators of women's status net of the variation accounted for by female age at marriage. A logical method for quantifying the 'net age difference' would be to include female age at marriage and female age at marriage squared as explanatory variables in addition to the women's status indicator in predicting the average age difference. However, there is a problem with this approach if there is measurement error in the variables used to derive the difference variable, and this can produce misleading results (e.g. Cain *et al*, 1992; Yanez *et al*, 1998). Chapter 3 calculates the average age difference as the difference in SMAMs so, because the SMAMs are based on census data rather than survey data, the extent to which the 'net age difference' would be subject to measurement error is questionable.

A problem does occur, however, when the age difference is used as a dependent variable in linear regression because the assumption that the error term has a mean of zero and constant variance does not hold (Plewis, 1985). Chapter 3 uses linear regression to obtain the adjusted R-square statistic to estimate the proportion of variation in the net age difference that can be accounted for by each indicator of women's status. However, this problem can be easily overcome since the average age at marriage of females appears on both sides of the equation, as an explanatory variable and as a component of the dependent variable, the age difference (Plewis, 1985). For example, if M denotes the male age at marriage and F denotes the female age at marriage, then the age difference is calculated as M-F, and this can be estimated as:

$$M - F = \alpha + (\beta - 1)\,F + \varepsilon$$

which can be re-arranged to:

$$M = \alpha + (\beta - 1 + 1)\,F + \varepsilon$$

which simplifies to:

$$M = \alpha + \beta F + \varepsilon$$

Chapter 3 uses this simplification to obtain the standardised residuals from regressing the male age at marriage on female age at marriage and female age at marriage squared. The residuals are then plotted against the women's status indicator to show graphically the bivariate association between each indicator of women's status and the net age difference. Standardised residuals are more useful than ordinary residuals as they have a mean of zero and a standard deviation of one, which enables the relative magnitudes of residuals to be compared.

2.3 Cross-national and within-country variations in the age difference using individual-level dat

A limitation of the analysis of WISTAT3 data is that within-country variations can not be studied since the data are national, aggregate estimates. Thus, Chapters 4, 5 and 6 proceed to use individual-level data from the Demographic and Health Survey (hereon referred to as the DHS) to examine both cross-national variations in the average age difference and variations in the distribution of the age difference within countries, for a sample of 30 less developed countries. This data source and the methods employed in Chapters 4, 5 and 6 are now discussed.

2.3.1 The Demographic and Health Survey

The World Fertility Survey (hereon referred to as WFS) was conducted from 1972 to 1984 and was the first attempt to collect nationally representative individual-level data from women in less developed countries on a number of demographic and health issues, such as fertility, maternal and child health, nuptiality, and household composition. In 1984, the DHS was introduced to continue the work of the WFS. The target population for the DHS is women, sometimes specifically ever-married women or those of childbearing age, which is usually defined as ages 15 to 49 inclusive. These women are identified through a household schedule that is administered to randomly selected households. Households are selected through a nationally representative probability sample, which is usually stratified at least by urban/rural area, and often by geographic or administrative area. Cluster sampling is then used to select primary sampling units within rural and urban strata. Some areas are over-sampled in an attempt to improve within-country estimates and comparisons, for example between urban and rural areas. A few surveys are designed to be self-weighting such as the 1993 DHS of Ghana. However, in general, it is necessary to weight the data to take account of variations in the probability of selection at the individual level. While the overall effect of the weighting is considered to be small (Lê and Verma, 1997), using weights ensures that more reliable population estimates are obtained.

At the time of writing, the DHS is in its third phase. Phase I was conducted between 1984 and 1989, phase II was conducted between 1988 and 1993, and phase III began in 1992. In terms of the questions asked in the DHS, there is an important distinction between the phases for the study of the spousal age difference. The woman's age at interview and age at the start of her first union have been asked in the core questionnaire of all phases. However, prior to phase III, her husband's age was only occasionally collected as part of additional modules. For the first time in phase III, the age of the woman's current husband or partner and also his line number in the household schedule were asked in the core questionnaire. It is possible therefore, at least in theory, to obtain two estimates of the husband's age: one as reported by the woman in the individual questionnaire, and one from linking data collected by the individual questionnaire to data collected by the household schedule about her husband.

Chapters 4 and 5 use DHS datasets with age data for both women and their husbands that were available for as many countries as possible at the time of data collation. This totals 30 countries. Table 2.3 shows that for most DHS surveys used in this study, the data come from phase III and thus typically refer to the mid-1990s, with the most recent interviews conducted in 1997 in Mozambique. There are six countries for which, at the time of data collation, phase III data were not available but for which both spouses' age data could be obtained from phase II. These are Burkina Faso, Cameroon, Malawi, Niger, Pakistan and Rwanda. The least recent data used in this study are from interviews conducted in Pakistan during 1990 and 1991.

Table 2.3 Target populations for the 30 DHS surveys used in Chapters 4 and 5 by UN region

Country by UN region	Year of fieldwork	Target population [1,2]	Number of women interviewed
Central & South America			
Brazil	1996	AW (de jure [3]) 15-49	12612
Colombia	1995	AW (de jure [3]) 15-49	11140
Dominican Republic	1996	AW 15-49	8422
Guatemala	1995	AW 15-49	12403
Haiti	1994/'95	AW 15-49	5356
Peru	1996	AW 15-49	28951
West Asia & North Africa			
Egypt	1995	EMW 15-49	14779
Kazakhstan	1995	AW 15-49	3771
Uzbekistan	1996	AW 15-49	4415
South Asia			
Bangladesh	1996/'97	EMW 10-49	9127
Nepal	1996	EMW 15-49	8429
Pakistan [4]	1990/'91	EMW 15-49	6611
Philippines	1993	AW 15-49	15029
Central, East & South Africa			
Central African Republic	1994/'95	AW 15-49	5884
Comoros	1996	AW 15-49	3050
Kenya	1993	AW 15-49	7540
Malawi [4]	1992	AW 15-49	4849
Mozambique	1997	AW 15-49	8779
Rwanda [4]	1992	AW 15-49	6551
Tanzania	1996	AW 15-49	8120
Uganda	1995	AW 15-49	7070
Zambia	1996	AW 15-49	8021
Zimbabwe	1994	AW 15-49	6128
West Africa			
Benin	1996	AW 15-49	5491
Burkina Faso [4]	1992/'93	AW 15-49	6354
Cameroon [4]	1991	AW 15-49	3871
Côte d'Ivoire	1994	AW 15-49	8099
Ghana	1993/'94	AW 15-49	4562
Mali	1995/'96	AW 15-49	9704
Niger [4]	1992	AW 15-49	6503

Notes for Table 2.3

1. AW = 'all women'
2. EMW = 'ever married women'
3. 'De jure' meaning 'by right', thus legally resident
4. The most recent data for the country are from phase II and not phase III

2.3.2 Selection criteria for the analysis of the DHS data

This is not the first study to use individual-level data to consider variations in the spousal age difference. As mentioned in Chapter 1, the study by Casterline *et al* (1986) used data from the WFS for this purpose. As their results refer to an earlier time period, Chapters 4 and 5 attempt to examine the extent to which their findings apply to more recent data. In order to permit comparisons between the two studies, it is therefore important to use the same sample selection criteria. Casterline *et al* (1986) used three criteria, which are that women should be:

1) currently married or living with partner,
2) in their first marriage or consensual union, and,
3) in a union that commenced within the ten years preceding the interview for the survey.

The first two criteria exclude women who have been divorced or widowed from the sample. This may bias the estimate of the average age difference, particularly in the case of widowhood since the deceased husband is likely to have been older, thus such age differences are likely to be generally wider. However, as in the WFS, only the age of the current husband or partner is obtained in almost all the DHS surveys.

Defining marriage as partners who live together rather than formal marriage has the advantage of including cohabiting couples in the sample. This is important since cohabitation is more common than marriage in some societies, especially in parts of Latin American and the Caribbean (Shryock and Siegel, 1976; Blanc and Rutenberg, 1990; Parrado and Tienda, 1997). Appendix 2C gives the proportion of women in each sample that was cohabiting rather than married at interview. Of the samples studied, there is much within and between region variation in the proportion of women who were cohabiting. Although, it is important to remember that a few DHS surveys only sampled ever-married women and so women who were unmarried but who had only ever cohabited or who were cohabiting at interview are excluded from these samples. Since the samples used in this analysis consist of both married and cohabiting couples, the terms 'husband' and 'wife' are used to refer to spouses for those in married unions as well as partners for those in cohabiting unions for simplicity and abbreviation.

A further problem regarding cohabitation is the subjectivity of its definition in terms of the frequency with which a couple must reside together in order to be considered as cohabiting (Singh and Samara, 1996). In attempt to overcome this, the DHS defines cohabitation as resident in the woman's household at interview, yet this introduces a further problem. If the husband's age is not recorded in the woman's questionnaire then he must be resident in the household if his age is to be recorded in the household schedule. However, it was estimated from surveys conducted for the World Health Organisation Global Programme on AIDS in five sub-Saharan African countries that between 11% and 43% of currently married men were not cohabiting with their wives, which it was suggested was due to seasonal or permanent migration (Caraël, 1995).

Defining couples as co-resident does not overcome the problem of determining when the union progressed from being a visiting-union to formal cohabitation, or the age of the partners at this time. Because of this ambiguity, some women may over-estimate the duration of a union, while others may ignore previous cohabiting unions especially if it occurred at a very young age or if it was of a short duration (Gage, 1995). This ambiguity is not confined to cohabitation. A number of studies have commented how it is difficult to define the beginning of a marriage since the ceremony is a process rather than a single event in some traditional societies (van de Walle, 1968; Ekpere *et al*, 1978; Vellenga, 1983; Lesthaeghe *et al*, 1989; UN, 1990; Meekers, 1992, 1996; Isiugo-Abanihe, 1994; Magnani *et al*, 1995). For example, Bloom and Reddy (1986) describe a traditional Indian wedding as a two-stage process in which the marriage ceremony and cohabitation are separated by a considerable time interval, the length of which varies according to the age of the bride.

Adopting a marriage cohort approach, the third criterion, has three advantages. First, this avoids the bias that a birth cohort analysis involves because some of the younger cohort members will not have experienced marriage or cohabitation by the time of interview. Second, there is evidence to suggest that data on the timing of recent unions are more likely to be of better quality (Rutstein and Bicego, 1990). Indeed, it is for this reason that using only data for women of childbearing age, i.e. the target population for the DHS surveys, is unlikely to affect the estimate of the age difference because only a negligible proportion of women begin their first union either later than age 49 or at a point within the age range 40 to 49 in the ten years preceding the survey so that they are then too old to be eligible for interview. Thirdly, given the correlation between age at marriage and the age difference, using a marriage-cohort approach is superior to looking at all the unions in a cross-sectional survey because in the latter, the unions of people from the younger birth cohort who have married at an early age are represented whereas those of people who will marry later are not.

2.3.3 Calculating the spousal age difference using individual-level data

In order to study the spousal age difference, there is the obvious criterion that age data are available for both spouses. Age is recorded by the DHS in completed-years for the wife's age at interview, the husband's age at interview, and the wife's age at marriage. The wife's birth date, marriage date, and date of interview are recorded as 'century-month-codes' (CMCs), which is the number of months since the beginning of the twentieth century. Potentially, this is a more accurate way of recording the wife's age at these events than expressing her age in completed years. Thus, expressing the age at marriage in completed-years under-estimates the exact age at marriage by 0.5 years on average, because for approximately half of the women, their age lies between .0 and .4999 when calculated from CMCs, while for the other half, their age lies between .5 and .9999. When averaged across the population it would be expected that the average age would be +0.5 years greater than when estimated from ages expressed in completed-years. The median age difference was obtained for all samples using both methods of estimating the wife's age. When the age difference is calculated using the wife's age from CMCs, it both under- and over-estimated the age difference when calculated from ages expressed in completed-years. For half of the 30 samples, the median age difference was 0.06 years smaller on average when calculated using the wife's age from CMCs, while for eleven samples, the median age difference was 0.05 years larger on average when calculated in this way. For the remaining four countries, there was no difference in the median age difference. Across all 30 samples, the average difference in the median age difference was negligible at -0.01 years.

While using a more accurate estimate of the wife's age is appealing for trying to obtain a more accurate estimate of the age difference, the impact on the age difference of *not* using a CMC-derived estimate of husband's age can also be considered. If the husband's age could be derived from CMCs, and the difference between this estimate and the completed-years estimate is assumed to be similar to the difference between the estimates of wife's age, then it is possible that the aggregate estimates of the age difference would not be very different to those obtained from using two ages expressed in completed-years because the values after the decimal point would, on average, cancel out. It is for this reason that this analysis calculates the age difference as the husband's age in completed years minus the wife's age in completed years. For consistency, age at marriage for both husbands and wives is calculated as the age at interview minus the number of years since the start of the union.

2.3.4 Quality of the individual-level age data

It is well documented that there is a tendency to misreport age particularly in traditional societies (van de Walle, 1968; Singh, 1985, 1987; Barbieri and Hertrich, 1999). This may be because little social significance is attached to chronological age, with rites of passage playing a greater role in this respect, as discussed in Chapter 1 (Parsons, 1940; 1942; Dorjahn, 1959; Brass *et al*, 1968). This section now investigates the quality of the age data in the DHS samples.

It has long been reported that age distributions heap on ages ending zero and five, and to some extent two and eight, while ages ending one and nine, and to some extent, three and seven are under-reported (Carrier, 1959; Brass *et al*, 1968; Nagi *et al*, 1973; Ewbank, 1981; Morah, 1985). Studies that doubt data quality sometimes try to measure the severity of the problem with an index of digit preference (e.g. Goldman *et al*, 1985; Morah, 1985; Singh, 1987; UN, 1987; Barbieri and Hertrich, 1999). Indices of digit preference are based on assumptions regarding the true age distribution over part or all of the age range and from this, comparisons can be made between the observed and expected frequencies. One such method is age ratio analysis, which compares observed and expected frequencies for individual ages (Shryock and Siegel, 1976). An age ratio is calculated as the ratio of the frequency for age x to the average of the frequencies for age x and the adjacent ages, age x-1 and age x+1. Age ratio analysis assumes that the true ages are approximately evenly distributed so that equal frequencies are expected for each age over a specified age range, such as three, five or ten years. Since this study is concerned with populations of less developed countries where high mortality and high fertility result in a pyramid-shaped age distribution, age ratio analysis is performed over a relative narrow age range of three years.

The age ratio is expressed as

$$\frac{F_x}{\frac{1}{3}(F_{x-1}+F_x+F_{x+1})} x100 = \text{Age ratio for age } x$$

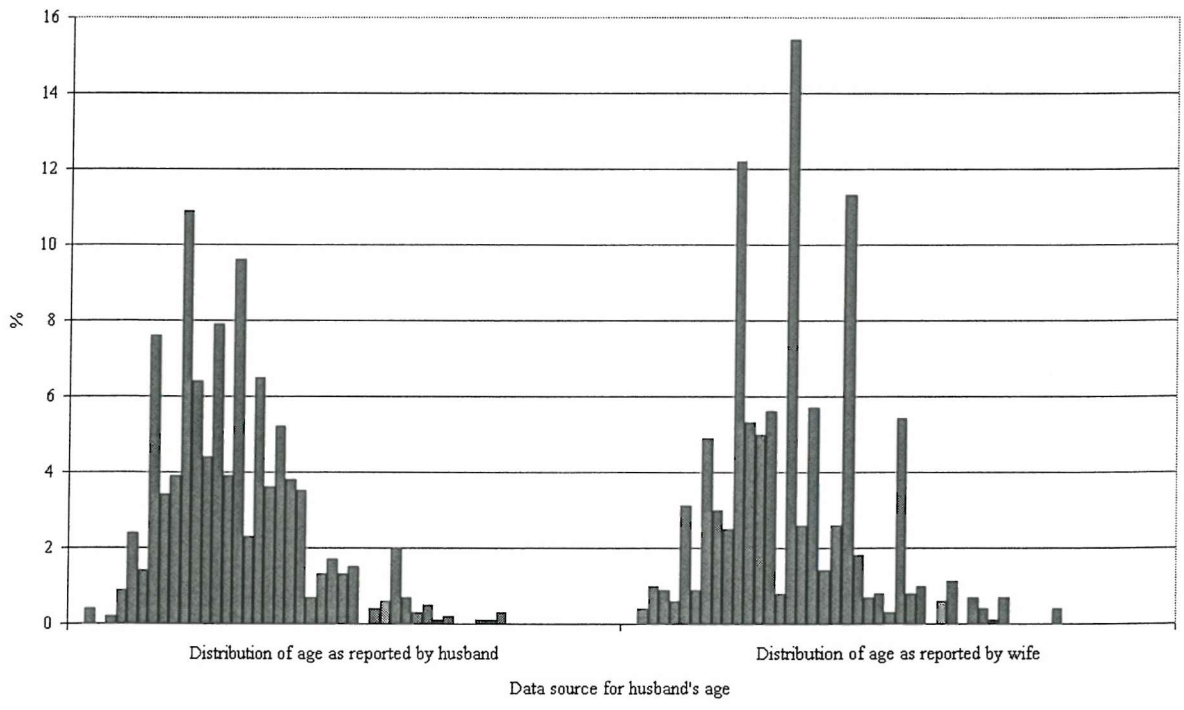
where F_x is the frequency of people aged x

Thus, age x is over-represented relative to its adjacent ages when the age ratio for age x is greater than 100. However, age ratios of more than 100 do not necessarily imply age misreporting. Fluctuations in F_x may also be due to fluctuations in births, changes in the age structure of those marrying as a result of variations in the tempo of marriage, or sampling variability (Shryock and Siegel, 1976). In an attempt to allow for these other sources of variation, this thesis considers ages with age ratios of more than 115 as over-reported, while ages with age ratios of less than 85 are considered as under-reported. While this definition of misreporting is based on arbitrary thresholds, it does allow for a reasonable amount of variation in the observed frequency of age x for reasons other than age misreporting.

In addition to identifying ages that are over- or under-reported (see Appendix 2D), the age ratio can be used to obtain a measure of the extent of age misreporting in a sample, i.e. which data source is likely to be more reliable overall, the woman's questionnaire or the household schedule. The 'mean absolute deviation' is calculated as a summary statistic for this purpose by summing the absolute deviation from 100 to ensure that none of the deviation is lost through age ratios of less than 100 cancelling out age ratios of more than 100, for ages in the range: median age $\pm$ five years. This age range avoids extreme values influencing the mean. Appendix 2E presents the mean absolute deviation for each data source for the wife's age and, where available, husband's age. It shows that, according to this measure, the woman's questionnaire provides a relatively more reliable estimate of the wife's age than the household schedule for almost all samples. Only three samples have a statistically significant smaller mean absolute deviation for the household schedule in comparison to the women's questionnaire, according to a t-test ($p<.05$). These are Kenya, Zambia and Malawi. Since this is a small proportion of samples, the woman's questionnaire is used as the data source for the wife's age for consistency and because it seems logical since this age is self-reported.

Of the 15 samples where the husband's age is available from both sources, the mean absolute deviation is smaller in the household schedule than in the woman's questionnaire for ten of these samples. Two possible reasons for this are that the wife's knowledge of her husband's age may be unreliable and/or that the husband may be the informant for the household schedule and so report his own age. This is particularly evident for the sample of husbands from Pakistan's DHS survey where the mean absolute deviation for husbands' ages from the woman's questionnaire is almost twice that estimated from the men's survey. Figure 2.1 compares the distribution of husband's age for the two data sources, and shows greater heaping in the distribution of ages reported in the women's questionnaire than in the distribution of ages reported in the men's survey.

Figure 2.1: Comparing the extent of age heaping using husband's age data as reported in the women's questionnaire with age data reported in the men's survey of the 1990/'91 Pakistan DHS survey



Source: 1990/'91 DHS of Pakistan

While it is possible that the husband was the informant for the household schedule and thus reported his own age, this is not certain and thus the household schedule is not automatically the preferred data source. Instead, where there is no statistically significant difference in the mean absolute deviations ($p>.05$) then the preferred data source for husband's age is that with the largest sample size. In most cases this is the woman's questionnaire because often only a fraction of husbands can be linked to the household schedule and their age obtained, as discussed above. For example, for the sample of Peru, over 7000 husband's ages are available from the woman's questionnaire in comparison to 504 from the household schedule.

In an attempt to summarise this study's data quality, samples can be categorised in terms of their mean absolute deviation. As discussed above, where the mean absolute deviation is less than 15 points then variations in the age ratios are not considered as problematic. However, where the mean absolute deviation is 15 or more points then variations in the age ratios are increasingly likely to be due to age misreporting. Appendix 2F ranks samples by their mean absolute deviation and shows that, according to these criteria, age misreporting is a problem in many of the samples, especially for the husband's age. Since the emphasis of this thesis is on the age difference, it is important to consider the impact of age misreporting on the estimates of the age difference. This is the objective of the next section.

2.3.5 Estimating the impact of age misreporting for the spousal age difference

When a man's age is over-reported, for example, a man who is 28 but reports his age as 30, then this over-estimates the couple's age difference, assuming that his wife is younger than him and that her age is accurately reported. Conversely, when a man's age is under-reported, for example, in the case of a man who is 42 but reports his age as 40, then this under-estimates the couple's age difference, again assuming that his wife is younger than him and that her age is accurately reported. For women, over-reporting age results is an under-estimation of the couple's age difference, assuming that she is younger than her husband and that his age is accurately reported. Conversely, when a woman under-reports her age then this over-estimates the couple's age difference, again assuming that she is younger than her husband and that her age is accurately reported. Of course, it is unknown whether for husbands and wives the misreporting is in the same direction, or whether it is of the same magnitude. Indeed, this assumes that it is possible to distinguish between individuals who misreport their age as age *m* and individuals who truly are aged *m*. The impact of age misreporting is therefore unknown for the age difference of individual couples. However, aggregate estimates of the impact of misreporting on the age difference are possible.

The mean spousal age difference for ages that are considered as misreported (Appendix 2D) is compared with the mean spousal age difference for the adjacent ages to see if there is statistically significant variation in the age difference. Of the 166 husbands' ages in the 30 samples that are considered as over- or under-reported, there is statistically significant variation in the mean age difference for just eight husbands' ages relative to the adjacent ages ($p<.05$ according to t-tests where variance homogeneity can be assumed, else the Mann-Whitney test). While of the 113 wives' ages in the 30 samples that are considered as over- or under-reported, there is statistically significant variation in the mean age difference for just four wives' ages relative to the adjacent ages. There are thus twelve ages in total for which there is statistically significant variation in the mean age difference, however, there is no statistically significant association in either the direction or the magnitude of this variation relative to the adjacent ages (since $p>.05$). Indeed, since significance is considered at $p<.05$, these twelve significant results are fully consistent with sampling error. To conclude, while it is likely that ages are over- and under-reported in many of the samples examined, age misreporting seems to have little impact on estimating the age difference.

2.3.6 Supplementing the age data

For some samples, the comparatively less reliable data source may have age data available for cases for which age data are missing in the more reliable data source. While it is possible to supplement the more reliable data source to increase the sample size, there is thus a trade-off between data quality and sample size. However, a degree of error is attached to all estimates, including those from the more reliable data source, and so it is considered sensible to use the best available estimate rather than omit couples from the analysis. This means that as many as 1,960 additional couples can be included in the analysis, as in the case of Pakistan.

Where the husband's age is only available from one data source and is missing, it is possible to link to the household schedule via the respondent's relationship to the household head. For example, if the woman is identified in the household schedule as the spouse of the household head, then her husband's age is the age of the household head. Similarly, if the woman is the household head, then her husband's age is the age of the household member whose relationship to the household head is recorded as spouse. It is possible to make assumptions about other relationships to the household head. For example, if the woman is the daughter of the household head then her husband will be recorded as the son-in-law. However, in some such households, there is more than one son-in-law so it is not possible to identify the correct son-in-law. In this situation assumptions could be made regarding the most likely age difference, but since this is the variable of interest, this approach was not considered sensible. Alternatively, women could be randomly allocated to a husband with the correct relationship to household head, but this assumes that her husband is resident in the household, which is not always the case, e.g. due to temporary migration (Schildkrout, 1983; Adepoju, 1983, 1994; Caraël, 1995; Hampshire and Randall, 2000). Furthermore, there are inconsistencies in the coding of the line number of husbands absent from the household (personal communication with Trevor Croft, Chief of Data Processing for the DHS, 2000).

Supplementing the more reliable data source results in a much higher median percentage of women eligible for this study for whom an estimate of their husband's age is available: 98.7% (lower quartile: 66.9% and upper quartile: 99.9%). This compares to a median of 81.5% (lower quartile: 29.5% and upper quartile: 99.8%) of women eligible for this study for whom an estimate of their husband's age is available when the data are not supplemented. Appendix 2G shows the proportion of women who meet the three sample selection criteria who have age data available for both spouses by supplementing data source.

2.3.7 Methods of analysis used in Chapters 4 and 5

2.3.7.1 Cross-national and within-country variations in the age difference using individual-level data

The statistical package SPSS (version 9) is used for the descriptive analysis of the individual-level age difference data. Descriptive statistics and graphs are used to illustrate variations in the age difference between and within samples. In the cross-national analysis, the median is used as the measure of location since the distributions are often highly skewed making the mean inappropriate. The median is expressed to two decimal places through extrapolation, and thus assumes an even distribution between an age difference of *n* years and *n*+1 years. The 95% confidence interval for the median is given in order to give an indication of its precision (Campbell and Gardner, 1988). To complement the median, the inter-quartile range is used as a measure of dispersion in the age difference within countries. In order to compare distributions in their entirety, the percentage distribution of age differences is categorised into five-year-groups: less than 0 years (i.e. the wife is older than her husband), 0-4 years, 5-9 years, 10-14 years, and at least 15 years. These categories are defined according to Casterline *et al* (1986) to enable comparisons with their analysis of the WFS data.

Given the tendency for polygamous unions to have larger age differences, on average (e.g. Dorjahn, 1959; Brown, 1981; Bongaarts *et al*, 1984; Garenne and van de Walle, 1989; Timaeus and Reynar, 1998), variations in the proportion of women married to a polygamous husband and also their rank relative to their co-wives are considered. Variations in the average age difference by these characteristics are shown graphically with boxplots. The statistical significance of these variations are examined using t-tests where there is variance homogeneity, else Mann-Whitney tests are used. Cross-national and within-sample variations in the age difference are then examined specifically for monogamous couples to permit more valid cross-sample comparisons. Although, it is important to acknowledge that union status refers to the time of the interview and so it is possible that a monogamous union becomes polygamous at a later date. This focus on the association between polygamy and the age difference was not reported by Casterline *et al* (1986).

Given that previous studies have reported an association between the average age difference and the average age at marriage for women cross-nationally (e.g. Bogue, 1969; Dixon, 1971; Nawar, 1988; Bozon, 1991; Ní Bhrolcháin, 1992), and an association between the age difference and the age at marriage of men within countries (e.g. Casterline *et al*, 1986, Ní Bhrolcháin, 1992), cross-national and within-country variations in the age at marriage of husbands and wives are also examined using the individual-level data. These variations are then repeated specifically for monogamous husbands and wives of monogamous men. As for the analysis of variations in the age difference, the medians, 95% confidence intervals for the medians, the inter-quartile ranges, and the percentage distributions are examined. The age at marriage of wives is systematically categorised for all samples as: before age 16, ages 16 to 18, ages 19 to 21, ages 22 to 24, and age 25 and older. While the age at marriage of husbands is systematically categorised for all samples as: before age 20, ages 20 to 22, ages 23 to 25, ages 26 to 29, and age 30 and older. Casterline *et al* (1986) did not report an examination of cross-national and within-country variations in the age at marriage of husbands and wives.

The associations between the age difference and the age at marriage of husbands and wives are then examined in turn. Scatterplots are used to show this association. The adjusted R-square is obtained from the regression of the age difference to estimate the proportion of the variation in the age difference that is explained by the age at marriage of husbands and wives, in turn. The adjusted R-square statistic is used in preference to the R-square statistic since the adjusted R-square takes into consideration the sample size as well as the number of explanatory variables.

2.3.7.2 Testing the hypotheses for variations in the age difference: Examining the 'marriage market hypothesis'

After examining the age difference distributions for 30 countries, Chapter 4 proceeds to use joint distribution theory to compare the observed distributions with the expected distributions of age differences. The expected distribution is that which would occur if unions were formed without any regard to spouse age, but instead was determined by the age-sex structure of eligible individuals in the marriage market.

In addition to line graphs to graphically display the observed and expected distributions, summary statistics similar to those presented for the observed distributions are obtained for the expected distributions. This approach replicates that used by Casterline *et al* (1986) to enable comparisons with their analysis of the WFS data.

2.3.7.3 Testing the hypotheses for variations in the age difference: Examining the 'by-product hypothesis'

Chapter 5 examines the hypothesis that the association between the husband's age at marriage and the wife's age at marriage at the individual level is a by-product of factors associated with the age at marriage of husbands and wives. Thus, this hypothesis assumes that the pattern of association between the husband's age and the wife's age is responsible for the systematic component of the distribution of age differences, i.e. the difference between the observed distribution and the distribution expected given the 'random' matching of spouses with respect to their ages at marriage. Chapter 5 begins by quantifying the association between the age at marriage of husbands and wives at the individual level using the Pearson product moment correlation coefficient for all samples of all couples, and also, just couples where the husband is monogamous for samples that collect data on polygamy. Casterline *et al* (1986) used the covariance for this purpose, however, the Pearson correlation coefficient is a standardised measure in relation to the variance, enabling comparisons to be made between samples. Scatterplots of the husband's age at marriage plotted against the wife's age at marriage are shown to illustrate these associations.

Casterline *et al* (1986) examined the by-product hypothesis by comparing the partial correlation coefficient with the Pearson correlation coefficient. The partial correlation coefficient is the correlation that remains between two variables when the linear effects of the other independent variables in the model have been removed from both (Norušis, 1992). Casterline *et al* (1986) calculated the partial correlation between the age at marriage of husbands and wives net of two variables: the number of years of schooling achieved by the husband and wife. However, the analysis in Chapter 5 calculates the partial correlation coefficient net of the linear effects of as many as eleven factors associated with age at marriage, depending on data availability. In addition, partial residuals are obtained from regressing the husband's age on the available factors, and similarly the partial residuals from regressing the wife's age on the same available factors. These residuals are then plotted to illustrate the net association between the ages at marriage of husbands and wives. The next section identifies the factors associated with age at marriage used in Chapter 5's analysis.

2.3.7.4 Factors that are hypothesised as associated with age at marriage

Fortes (1978) argued that understanding what governs the timing of marriage in a population requires consideration of the social and geographical range within which partners are found, which in turn, depends on the rules of permitted and prohibited mating such as lineage, exogamy and cousin marriage. As discussed in Chapter 1, a number of studies have identified evidence of an association between a variety of factors and age at marriage. This thesis draws on this evidence in its choice of factors for this analysis, although this choice is limited by the data available from each DHS survey. Unlike the DHS, the Asian Marriage Survey of the late 1970s asked questions specifically concerned with union formation such as the relationship between families before marriage, the type of dowry, and the wealth flows at marriage (Fricke *et al*, 1986). While it is desirable to have such specific information for the examination of the by-product hypothesis, it is only appropriate to ask such questions in particular societies. For example, while all DHS surveys ask about the highest level of education achieved by both husband and wife, not all DHS surveys ask about consanguinity (whether the husband and wife are blood relatives), or whether the husband is polygamous. A balance must be struck therefore between data availability and the ability to compare each country's results. In addition, it is important to acknowledge that the associations between the ages at marriage of husbands and wives and the factors associated with age at marriage may vary if considered in different social and economic contexts (Jones, 1980).

This analysis considers eleven factors that have an association with age at marriage at the individual-level. Table 2.4 shows the availability of data for these factors by country. All the factors hypothesised as associated with age at marriage are categorical variables. Some of the factors are available directly from most DHS surveys. These factors and their categories are:

place of residence: urban or rural; union status: married or cohabiting; wife's highest level of education: none, primary, secondary or higher; husband's highest level of education: none, primary, secondary or higher. In contrast, some of the factors are derived from more than one variable. These factors and their categories are: household socio-economic status: low, average, or high; the wife works and earns cash: yes or no; household structure: nuclear, extended, or polygamous; consanguineous union: first cousin, second cousin/other relation, or not related; polygamous union: monogamous, polygamous and wife is rank one, or polygamous and wife is of rank two or higher. Appendices 2H and 2I give the sample percentage distribution for each factor, by country.

Table 2.4 Data availability for eleven factors that are hypothesised to be associated with age at marriage, by country and UN region

Country by UN region	Factors hypothesised to be associated with age at marriage:										
	Place of residence	Married/cohabiting	Household S.E.S.	Wife works & earns cash	Wife's education	Husband's education	Religion	Ethnicity	Household structure	Consanguinity	Polygamous husband
Central & South America											
Brazil	•	•	•	•	•	•	•	•			
Colombia	•	•	•	•	•	•					
Dominican Republic	•	•	•	•	•	•					
Guatemala	•	•	•	•	•	•	•	•			
Haiti	•	•	•	•	•	•	•		•		
Peru	•	•	•	•	•	•					
West Asia & North Africa											
Egypt	•		•	•	•	•	•			•	
Kazakhstan	•	•	•	•	•	•	•				
Uzbekistan	•		•	•	•	•	•	•			
South Asia											
Bangladesh	•		•	•	•	•	•		•		
Nepal	•		•	•	•	•	•				•
Pakistan	•		•	•	•	•				•	•
Philippines	•	•	•	•	•	•	•	•			
Central, East & South Africa											
Central African Republic	•	•	•	•	•	•	•		•	•	•
Comoros	•	•	•	•	•	•			•	•	•
Kenya	•	•	•	•	•	•	•		•		•
Malawi	•	•	•	•	•	•			•		•
Mozambique	•	•	•	•	•	•	•		•		•
Rwanda	•	•	•	•	•	•	•	•			•
Tanzania	•	•	•	•	•	•			•		•
Uganda	•	•	•	•	•	•	•		•		•
Zambia	•		•	•	•	•	•		•		•
Zimbabwe	•		•	•	•	•	•	•	•		•
West Africa											
Benin	•	•	•	•	•	•	•		•		•
Burkina Faso	•		•	•	•	•	•		•		•
Cameroon	•	•	•	•	•	•	•	•	•		•
Côte d'Ivoire	•		•		•	•	•		•		•
Ghana	•	•	•	•	•	•	•				•
Mali	•	•	•	•	•	•	•		•		•
Niger	•		•	•	•	•	•	•	•		•

There is the issue of data sparsity, which can result in model instability (Edwards, 1995). In an attempt to minimise the number of dimensions in the data, and thus how sparsely the data are distributed across the variable categories, categories are sometimes collapsed. For example, where less than 5% of women have more than secondary education, these women are categorised with women who have secondary education. The method of collapsing categories is determined by the distribution of each sample, which does vary between samples although similarities do emerge. The greatest number of categories observed for any factor is five (for religion in the Central African Republic), but generally variables tend to have two or three categories. In some instances, such as ethnicity, it is substantively inappropriate to combine categories (Fapohunda, 1983; Timaeus and Reynar, 1998; Hampshire and Randall, 2000). For example, Isiugo-Abanihe (1994, p. 7) described Nigeria as *"ethnically complex with as many as 374 ethnic groups"*. For samples with more than five ethnicity categories, such as the 1994 Côte d'Ivoire DHS survey, which collected data on 14 ethnic groups, the variable is excluded from the analysis in order to avoid misleading or meaningless results. Appendix 2I gives details of how categories are collapsed for the education, religion and ethnicity factors and the resulting percentage distribution by country.

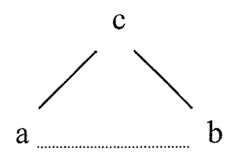
Relative to the other DHS surveys, the 1995 Egyptian DHS survey is a particularly rich data source. In addition to the women's individual questionnaire common to all DHS surveys, the 1995 Egyptian DHS survey included a module on the status of women that asked about the timing and circumstances of their marriage. At the time of data collation, this was the only DHS with this additional module, but since then six other DHS surveys have been conducted with this additional module: Tanzania, the Dominican Republic (both in 1999), Cambodia, Columbia, Peru, and Haiti (all in 2000). One possible limitation of using these data is that only half of the women eligible for the 1995 Egyptian DHS were invited to complete the women's status questionnaire. However, since these women were randomly selected, it is assumed that selection bias is minimal. The sample used for the analysis in Chapter 5 is therefore a sub-sample of the women used for the analysis in Chapter 4.

Five factors are used from the women's status module of the 1995 Egyptian DHS survey to complement, and in some cases, to supersede the factors available from the women's individual questionnaire. These factors and their categories are: wife's childhood place of residence: urban or rural; residence when first married: urban or rural; wife worked before marriage: yes or no; wife knew and chose husband: wife knew and chose husband herself, wife knew but family/others chose husband, or wife had not met her husband before marriage; husband's profession, grouped: professional/managerial, sales and service, skilled/unskilled production, agriculture, none.

2.3.7.5 Graphical modelling

A limitation of using the partial correlation coefficient and partial residuals is that these methods rely on pairwise marginal associations, which can result in identifying spurious associations (Edwards, 1995). For example, consider three variables, a, the husband's age at marriage; b, the wife's age at marriage; and c, a factor associated with the age at marriage of both husbands and wives, such as the level of education achieved by the husband.

Figure 2.2 Hypothesised association between the age at marriage of husbands and wives and husband's education



The solid lines in Figure 2.2 denote that the husband's education, c, is directly associated with both the age at marriage of husbands, a, and wives, b. However, the dotted line in Figure 2.2 denotes that the association between a and b exists because of c, in other words, it is a spurious association. This is an inadequacy of only studying marginal associations, rather than both marginal and conditional associations. Consequently, it is necessary to take a multivariate approach that includes all the relevant variables in the analysis simultaneously in order to study the conditional as well as marginal associations. This is the basis of graphical modelling. Chapter 5 uses graphical modelling to model the joint distribution of the age at marriage of husbands and wives net of the factors associated with age at marriage as identified in Table 2.4. It is therefore possible to ascertain whether two variables are conditionally independent given all other variables, or whether there is a direct association between the two variables (Mohamed *et al*, 1998). Thus, in the context of this study, it is possible to ascertain whether there is a direct association between the two ages at marriage, or whether they are conditionally independent given, for example, the wife's education, religion, or other factors associated with age at marriage.

The graphical modelling package MIM3.1, which is an acronym for 'Mixed Interaction Modelling, version 3.1' is used as it was designed for the hierarchical interaction mixed modelling of both discrete and continuous variables, as its name suggests (Edwards, 1995). However, there is a limitation involved with using MIM3.1 regarding the size of the data matrix. Although the number of factors and the number of categories within the factors is limited, it is still necessary to take a random sample of couples from some of the larger samples to reduce the sample size to one that is manageable for analysis in MIM3.1. Appendix 2J gives the reduced sample sizes used for the graphical modelling in MIM3.1.

In MIM3.1, the variables are considered in two groups or 'blocks'. The first block consists of the explanatory variables, i.e. the factors associated with age at marriage identified in Table 2.4. The second block consists of the outcome variables, i.e. the two ages at marriage. Starting with the saturated model, in which all the variables are mutually dependent, backwards stepwise selection is used to identify a parsimonious model. The statistical significance of each 'edge', that is, an association between two variables, is tested with an edge exclusion test that is based on the F-statistic, where statistical significance is considered at $p<.05$. At each step, the least statistically significant edge is removed. MIM3.1 uses the principal of coherence such that once an edge between two variables is identified as statistically significant, it remains in the model and its significance is not re-tested (Edwards, 1995). Backwards selection is preferred to forward selection for substantive reasons. Backwards selection starts with the saturated model and proceeds to refine this model, while forward selection begins with a model that initially consists of just the constant term, which is not a logical starting point (Edwards, 1995). First, the statistical significance of associations between the variables in block one are considered, then the associations between the variables in blocks one and two given the associations identified between the variables in block one, and finally, the association between the variables in block two, given the preceding associations. None of the edges are fixed, which means that associations are not forced between variables. Consequently, it is possible to see if there remains a statistically significant association between the age at marriage of husbands and wives net of the factors associated with age at marriage. This association is tested net of all the possible interactions between the variables in the model so it can be considered as the most thorough test of this association. The conditional independence structure between each pair of variables is presented in the form of a graphical model that provides clear visual representations of the relationships between variables. In addition, partial residuals are obtained from modelling each age at marriage net of the marginal and conditional associations with all other variables. The partial residuals for the husband's age at marriage are plotted against the partial residuals for the wife's age at marriage to show the net association, which is measured with the partial correlation coefficient. Unlike partial correlation coefficients obtained from linear regression, this partial correlation coefficient takes account of both marginal and also conditional associations with the other variables.

2.4 Examining the hypothesis that the age difference reflects the status of women at the individual level

Chapter 6 examines the hypothesis that the age difference reflects the status of women using individual-level data for one particular country, Egypt. While this hypothesis is examined in Chapter 3, it can not be assumed that the conclusions drawn from the aggregate-level analysis apply at the individual-level. As mentioned above, at the time of data collation, the 1995 Egyptian DHS survey was the only survey to include direct questions on the status of women, and thus, it is for this reason that Egypt is focussed upon in Chapter 6. Despite this data limitation, Egypt is an interesting country to focus this analysis upon. It is the most populous Arab country (Obermeyer, 1992), with a predominantly Muslim population (e.g. Central Intelligence Agency, 1999). As discussed with respect to the WISTAT3 analysis, the Islam faith has important implications for the status of women. One argument is that *"both religious law and cultural practices severely restrict women's non-reproductive options and their freedom of movement at the same time that* *they encourage and reward motherhood"* (Govindasamy and Malhotra, 1996, p. 329). However, others disagree with this stance -not least because it overlooks the complexities and variations in women's status within- and between- Islamic societies (Balk, 1994). Indeed, at the other extreme, there is the argument that Islamic law emphasises egality, including with respect to human rights and reproductive decision-making. For example, it has been argued that, because of the strict division of gender roles that are considered crucial to the success of the family, and which are often associated with Islam, women are likely to have substantial control over family matters (Govindasamy and Malhotra, 1996).

2.4.1 Hypothesised indicators of women's status from the 1995 Egyptian DHS survey

The women's status module of the 1995 Egyptian DHS asked women questions about a number of different aspects of women's status. This included questions on perceptions such as the relative costs of sons and daughters, and attitudes towards gender roles, as well as actual behaviour such as decision-making and financial empowerment. Given the numerous variables which may be considered as indicators of women's status, and the earlier discussion regarding how women's status is conceptualised and measured (section 2.2.4), it was considered necessary to focus this analysis.

Chapter 6 takes a lead from a DHS Occasional Paper titled *'Autonomy and Egyptian women'* published in 1995 by Sunita Kishor, one of the authors of the 1995 Egyptian DHS Final Report, who also co-wrote the DHS comparative study of the status of women in 25 countries (Kishor and Neitzel, 1996). Kishor's (1995) occasional paper used data from the 1988 Egyptian DHS since the core questionnaire included a number of questions on the status of women, the majority of which were incorporated in the women's status module of the 1995 Egyptian DHS. The status of women was considered specifically in terms of women's autonomy, and its association with contraceptive use and child mortality. In an attempt to consider autonomy multidimensionally, three composite indices of women's autonomy were derived from a number of the questions on women's status. Other studies that have derived indices of women's status for this reason include Balk (1994, 1997) and Schuler *et al* (1997), all in the context of Bangladesh; Niraula and Morgan (1996) in the context of Nepal; and Bloom *et al* (2001) in the context of India. In the context of Kishor's (1995) study, the indices attempt to differentiate between who should have decision-making powers within the household and who actually does have this power (Mason, 1995). There are two indices that try to capture perceived autonomy to distinguish between decision-making in relation to childbearing and child rearing and decision-making outside of this arena, which are termed, respectively, the 'customary autonomy index' and 'the non-customary autonomy'. The 'realised autonomy index', as its name suggests, tries to capture the autonomy that women actually have. Since devising indicators of women's status is not a principal objective of this thesis, and given that Kishor's (1995) indices are specific both in terms of the cultural setting, i.e. Egypt, and the data available, i.e. the DHS, Chapter 6 proceeds to derive these indices using data from the 1995 Egyptian DHS survey.

2.4.2 Deriving indices of women's autonomy using Kishor's (1995) method

Scores of either zero or one are assigned to the responses to questions that seek to reflect particular dimensions of women's autonomy. For example, responses to the question *"who has the final say regarding the children's education?"* (El-Zanaty *et al*, 1996, p. 332) are used to create the customary autonomy index. In this instance, women who say that either they have the final say, or have it jointly with their husband score one, while women who say either their husband or someone else has it score zero. Kishor (1995) justified this assignment of scores rather than distinguishing between women who have sole decision-making power and those who do not with the argument that, because Egypt is a patriarchal society, women believing that they should have decision-making powers at least equal to their husband is a significantly strong indication of women's autonomy, especially since the questions concern the family rather than specifically the woman.

Since the 1995 women's status module builds on the questions asked in the 1988 DHS, there are slight variations in how the indices are defined for this analysis relative to Kishor's (1995) study. Appendix 2K compares the questions used to derive the indices in Chapter 6 with the questions used by Kishor (1995). Table 2.5 shows the scores assigned to the questions used to derive each index in this study. As in Kishor's (1995) study, the equally weighted scores are then summed for the five variables that constitute each index, and the distributions are approximately evenly split into three levels of autonomy: low, average, and high.

Table 2.5 Scores assigned to questions used to derive 3 indices of women's autonomy

Index and components	Response and score			
Customary Autonomy Index "Who has the final say in your family on the following ...?"	Respondent	Husband	Both	Other(s)
Having another child	1	0	1	0
Children's education	1	0	1	0
Children's marriage plans	1	0	1	0
Medical attention for the children	1	0	1	0
Use of family planning methods	1	0	1	0
Non-Customary Autonomy Index "Who has the final say in your family on the following ...?"	Respondent	Husband	Both	Other(s)
Visits to friends and family	1	0	1	0
Household budget	1	0	1	0
"Agree/disagree..."	Agree	Disagree	Don't know	
A wife's place is not only in the home	1	0	0	
If a wife disagrees with her husband, she should speak-up	1	0	0	
"Do you know how much your family's approximate total income from all sources is?"	Yes 1	No 0		
Realised Autonomy Index "Are you usually allowed to got o the following places...?"	Alone	Only with children	Only with another adult	Not at all
Local market to buy things	1	0	0	0
Local health centre/doctors	1	0	0	0
Neighbourhood for recreation	1	0	0	0
Home of friends/relatives	1	0	0	0
"Who mainly decides how your family's income is spent?"	Respondent 1	Husband 0	Both 1	Other(s) 0

Source: 'Questions used and weights given to each possible response when constructing the autonomy indices', p. 4, in Kishor, S. (1995) *Autonomy and Egyptian women: Findings from the 1988 Egypt Demographic and Health Survey*. Demographic and Health Surveys Occasional Paper no. 2. Calverton, Maryland: Macro International Inc.

2.4.3 Sample selection criteria for the analysis of data from the 1995 Egyptian DHS

This analysis adopts a marriage cohort approach for the same reasons discussed in the context of the main DHS analysis (see section 2.2.2). In addition, restricting this analysis to women married for the first time in the ten years preceding the DHS interview allows a contemporary picture of women's autonomy in Egypt to be presented. The analysis makes use of the data collected from the additional module on women's status, which was asked of a random sample of half the women eligible for the individual questionnaire, which for this DHS is defined as ever-married women aged 16-44 at interview. Furthermore, since the women's autonomy indicators distinguish between traditional roles, i.e. considering women as wives and mothers, and non-traditional roles, i.e. roles outside of this sphere, 99 women without children are excluded from this analysis so that all women are eligible for all of the autonomy indicators. Of the sample of 5,352 ever-married women interviewed for the 1995 Egyptian DHS survey and who met the sample selection criteria, 1102 (1168 weighted) women completed the women's status module and had children, and are thus eligible for this analysis.

2.4.4 Methods of analysis used in Chapter 6

Chapter 6 begins by examining the percentage distribution of responses used to derive the three indices of women's autonomy as shown in Table 2.5. In an attempt to consider the extent to which the variables used to derive the indices are capturing different elements, the percentage of women who consistently report a less autonomous response and the percentage of women who consistently report a more autonomous response is compared for each of the indices. In addition, for each index Cronbach's Alpha is calculated as an estimate of reliability based on the correlations between the five variables that constitute the index (Norušis, 1992).

The variation in the average age difference is then considered for each of the 15 variables in order to consider women's autonomy unidimensionally. The percentage distribution of the summed scores for each index is then examined. For each index, variations in the average age difference are considered by autonomy level. Given the evidence in the literature review for an association between the age difference and the age at marriage of men within countries (e.g. Casterline *et al*, 1986, Ní Bhrolcháin, 1992), and the association between the age at marriage of women and the status of women at the individual level (e.g. Mason, 1985; 1986), Chapter 6 examines the variations in the average age at marriage of husbands and, in turn, wives by level of autonomy for each index.

The association between the age difference and each autonomy index is then examined net of the association between the age difference and in turn, the husband's age at marriage and then the wife's age at marriage. As discussed in section 2.2.6, to overcome the problem of logical dependency that exists when either the husband's or wife's age at marriage is controlled for by including it as an explanatory variable when the dependent variable is the age difference (Yanez *et al*, 1998), the graphical modelling package, MIM3.1 is used in Chapter 6 to model the joint distribution of the age at marriage of husbands and wives. In the first hierarchical interaction model, the first block of variables – the explanatory variables – consists of just the three autonomy indices. Then, the model takes into account the association between these autonomy indices and the age at marriage of husbands and wives net of the associations with the factors that are identified as associated with age at marriage in Chapter 5. In both models, the partial residuals are obtained for husbands and wives, and plotted against one another to see whether or not an association remains between the two ages. This is quantified using the partial correlation coefficient. A net association between one of the autonomy indices and one spouse's age at marriage is equivalent to an association between that autonomy index and the age difference net of the other spouse's age.

2.5 Conclusions

This chapter has reviewed some of the data issues involved in examining the hypotheses described in Chapter 1. By discussing the methods employed to achieve the results presented in Chapters 3 through 6, it is hoped that this has provided an overview of the aggregate to individual-level approach adopted by this thesis. The next chapter, Chapter 3, presents the results of the aggregate-level analysis of variations in the average age difference using data from the WISTAT3 database.

Chapter 3:
Results of examining the hypothesis that the spousal age difference reflects women's status at the aggregate-level

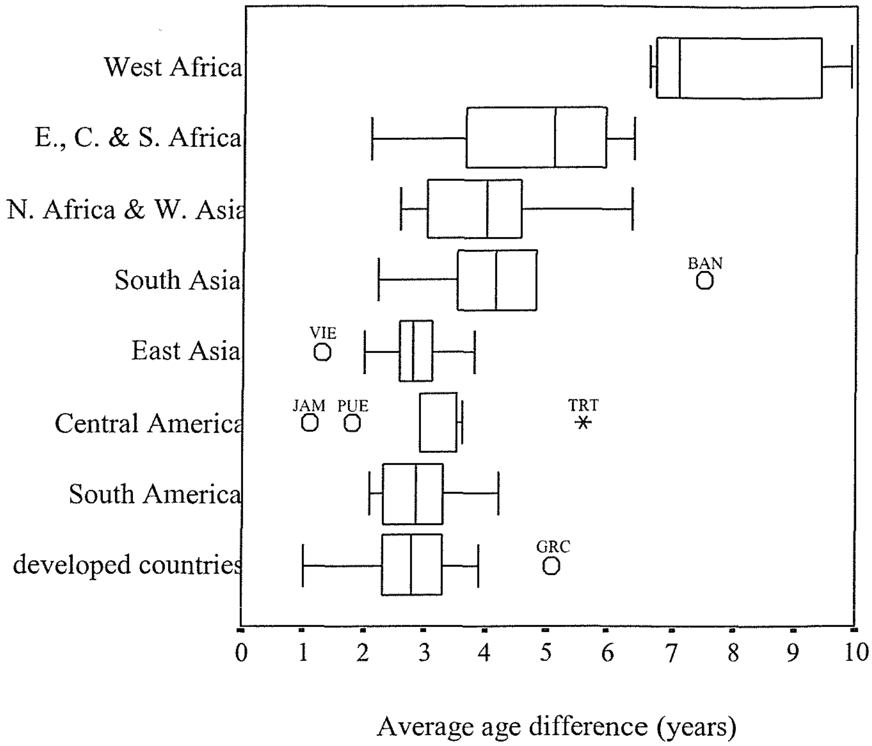
3.1 Overview

This chapter describes variations in the average age difference between and within United Nations' regions using aggregate-level data for 93 countries from the third edition of the United Nations' Women's Indicators and Statistics database. The age difference is measured as the difference between the male Singulate Mean Age at Marriage ('SMAM') and the female SMAM. Variations in the average age at marriage are also examined, as is the association between the age difference and its constituent parts. This chapter then proceeds to examine the hypothesis that the age difference reflects women's status at the aggregate level. This is achieved by examining the association between the average age difference and several hypothesised indicators of women's status, which this analysis considers in several different ways to reflect the ambiguity that surrounds how women's status is conceptualised and measured. Finally, this association is then considered net of female age at marriage since this is considered as a possible confounding effect.

3.2 Regional variations in the average age difference

Figure 3.1 shows boxplots of the distributions of age differences as calculated as the male SMAM minus the female SMAM, by UN region. As discussed in Chapter 2, it is important to remember that this sample of countries is incomplete and that some regions are better represented than others. Table 3.1 shows some summary statistics and the number of countries included in this sample by UN region. While the median age difference is positive for all regions, it is evident that there is much between and within region variation. This regional variation is highly statistically significant (p=.000 according to an analysis of variance) with more of the variation in the average age difference occurring between regions (58.3%) than within regions (41.7%).

Figure 3.1 Boxplots showing variations in the average age difference by UN region



Notes for Figure 3.1
1. E., C. & S. Africa denotes 'East, Central and South Africa'
2. N. Africa & W. Asia denotes 'North Africa and West Asia'
3. O denotes an outlying value
4. * denotes an extreme value
5. BAN = Bangladesh, VIE = Viet Nam, JAM = Jamaica, PUE = Puerto Rica, GRC= Greece, TRT = Trinidad and Tobago
6. See Appendix 2A for list of countries in each region

Table 3.1 Summary statistics for variations in the average age difference by UN region

UN region	Median	Inter-quartile range	Minimum	Maximum	n
West Africa	7.1	2.9	6.6	9.9	6
East, Central and South Africa	5.1	3.1	2.1	6.4	12
North Africa and West Asia	4.0	1.8	2.6	6.4	12
South Asia	4.2	2.3	2.2	7.5	6
East Asia	2.8	0.6	1.3	3.8	11
Central America	3.5	1.2	1.1	5.6	9
South America	2.9	1.1	2.1	4.2	10
Developed countries	2.8	1.1	1.0	5.0	27

The most striking feature of Figure 3.1 and Table 3.1 is the magnitude of the average age difference among the West African countries relative to countries in the other regions. The median age difference for the West African sample is 7.1 years. It is clear from the boxplot that its distribution is highly skewed, yet the smallest average age difference observed for this region is still 6.6 years for Benin, which exceeds all the other regional medians. While other studies have also observed significantly larger average age differences in West Africa (e.g. Timaeus and Reynar, 1998), such studies have often included second and higher order polygamous marriages in their sample (e.g. Casterline *et al*, 1986), for whom it is well-established that age differences tend to be larger, on average, than for monogamous couples and first polygamous marriages (e.g. Goldman and Pebley, 1989; Lesthaeghe *et al,* 1989). However, since the age difference was measured as the age difference at first marriage for both men and women, it is interesting to observe that average age differences also tend to be larger for couples in their first marriage in West Africa, and to some extent in the rest of sub-Saharan Africa, relative to other regions. There remains statistically significant regional variation in the average age difference if the West African countries are excluded from the analysis (p=.000 according to an analysis of variance). However, most of the variation occurs within regions when the West African countries are excluded, with less than one third of the total variation in the average age difference due to variations between regional samples.

There is only one other country in this sample which has a median age difference of the magnitude observed for the West African countries and that is Bangladesh at 7.5 years. Other studies that have used data from different sources have also observed an average age difference of this magnitude for Bangladesh (e.g. Stoeckel and Chowdhury, 1984; Casterline *et al*, 1986; Balk, 1994). These results suggest that the median for Bangladesh is unique when compared to the median for its regional sample, 4.2 years, even though the sample of South Asian countries used in this analysis is incomplete. This has also been observed by other studies (e.g. Dixon, 1971; Casterline *et al,* 1986; Bittles *et al,* 1993; Niraula and Morgan, 1996). In looking for explanations, one hypothesis that can be ruled out is that the large average age difference observed in Bangladesh is due to widows remarrying because traditionally in this society women do not remarry when widowed (Goldman and Pebley, 1989).

Some of the other countries in sub-Saharan Africa have average age differences that approach the magnitude of average age differences in the West African sample, such as Somalia at 6.4 years and Ethiopia at 6.2 years. However, there are also countries in this regional sample where the median is much smaller, less than half these estimates, such as for Swaziland at 2.4 years and 2.1 years for Reunion. This is also true of the sample of North African and West Asian countries. For example, Oman has a median age difference of 6.4 years, which compares to the smallest median observed for the North African and West Asian countries, 2.6 years for the United Arab Emirates. For the remaining UN regions, East Asia, Central and South America, and the developed counties, Figure 3.1 shows that there is still much variation within these regions but the inter-quartile ranges tend to be smaller at around one year or less. The sample of Central American countries is interesting in this respect. Although small (n=9), a third of the sample is identified as either an outlying or an extreme value (Jamaica, Puerto Rico and Trinidad and Tobago.) Thus, the difference between the largest and smallest age differences is 4.5 years, which is more than three times the inter-quartile range for this regional sample. Age differences tend to be smallest in the developed countries. The median for this sample of developed countries is 2.8 years, and values range from 1.0 year (Ireland) to 5.0 years (Greece). Although as Figure 3.1 shows, Greece is an outlier relative to the other developed countries, as the next largest average age difference for the developed countries is 3.9 years in Italy and the former Yugoslavia.

One criticism of calculating the age difference as the difference in SMAMs is that by definition, it refers only to the average age difference among married couples. However, in some societies, especially in Central and South America, the prevalence of cohabiting and visiting unions is relatively large (e.g. Shyrock and Siegel, 1976; Blanc and Rutenberg, 1990; Parrado and Tienda, 1997). Focussing solely on married couples therefore may ignore a large proportion of the population in union. This may in part explain the large within-region variation observed, especially for the sample of Central American countries.

3.3 Regional variations in the average age at marriage

Table 3.2 shows summary statistics for the average age at first marriage for females and males, which are both expressed as the SMAM, and the number of countries included in this sample by UN region. Figure 3.2 follows and shows boxplots of the distributions of the average age at first marriage for females and males by UN region.

Table 3.2 Summary statistics for variations in the average age at first marriage by UN region

Females

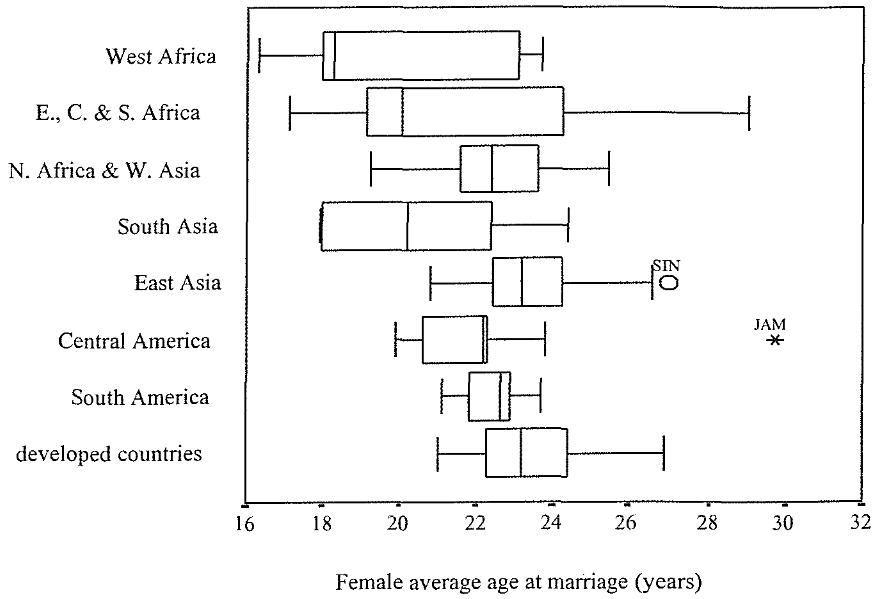
UN region	Median	Inter-quartile range	Minimum	Maximum	n
West Africa	18.3	5.7	16.3	23.7	6
East, Central and South Africa	20.1	6.0	17.1	29.0	12
North Africa and West Asia	22.4	2.1	19.2	25.5	12
South Asia	20.2	4.9	17.9	24.4	6
East Asia	23.2	2.3	20.8	27.0	11
Central America	22.2	2.5	19.9	29.7	9
South America	22.7	1.4	21.1	23.7	10
Developed countries	23.2	2.3	21.0	26.9	27

Males

UN region	Median	Inter-quartile range	Minimum	Maximum	n
West Africa	27.8	5.4	23.7	30.4	6
East, Central and South Africa	26.2	3.5	22.7	31.4	12
North Africa and West Asia	26.2	1.9	25.2	28.8	12
South Asia	25.1	3.9	21.5	27.9	6
East Asia	25.3	3.1	24.5	29.8	11
Central America	25.1	3.8	23.5	30.8	9
South America	25.5	0.8	24.3	26.0	10
Developed countries	26.2	2.6	24.2	30.3	27

Figure 3.2 Boxplots showing variations in the average age at first marriage by UN region

Females



Males

Notes for Figure 3.2

1. E., C. & S. Africa denotes 'East, Central and South Africa'
2. N. Africa & W. Asia denotes 'North Africa and West Asia'
3. O denotes an outlying value
4. * denotes an extreme value
5. SIN = Singapore, JAM = Jamaica
6. See Appendix 2A for list of countries in each region

Unlike the boxplot showing variations in the average age difference by region (Figure 3.1), Figure 3.2 shows that all of the regional inter-quartile ranges for the average age at first marriage for females and males overlap. However, it is evident from Figure 3.2 and Table 3.2 that there is more regional variation for the average age at first marriage for females than for males. This regional variation in age at first marriage is statistically significant for females but not for males (p=.020 and p=.403, respectively, according to Kruskal-Wallis tests since variance homogeneity can not be assumed). For both males and females, a greater proportion of the total variation in the average age at marriage occurs within region rather than between regions. For females, 80.0% of the total variation in female average age at marriage occurs within regions, while for men, almost all of the variation in male average age at marriage occurs within regions, 91.3%. For example, the regional sample with the largest inter-quartile range and range is East, Central and South Africa, with the average age at first marriage for females ranging from 17.1 years in Ethiopia, to 29.0 years in Swaziland, a range of 11.9 years. Relative to the total sample of 93 countries, these are the second youngest and second oldest average ages at marriage for females, respectively. The youngest female average age at marriage in this sample is observed as 16.3 years for Niger, which is 2 years below the regional average for this sample of West African countries. The inter-quartile range for the sample of West African sample is of a similar magnitude to that observed for the East, Central and South African countries at 5.7 years. Also, like the distribution of average age differences for the West African countries (Figure 3.1), Figure 3.2 shows that the West African distribution of average age at marriage for females is highly skewed, with the average age at marriage for females as great as 23.7 years in Senegal. The oldest average age at first marriage for females of the 93 countries studied here is observed as 29.7 years for Jamaica. Figure 3.2 highlights Jamaica as an extreme case among the sample of Central American countries, that is, it has a value that is greater than 3 times the inter-quartile range of 2.5 years from the upper quartile. Excluding Jamaica, the oldest average age at marriage for this regional sample is 23.8 years for Haiti, which is 6 years younger than for Jamaica.

There is little variation in the average age at first marriage for males between regions. There is 2.7 years difference between the youngest regional median (observed for the South Asian and Central American samples) and the oldest regional median (observed for the West African sample). Thus, as identified above, there is more variation in the average age at first marriage for men within regions rather than between regions, with the largest within-region inter-quartile range in male age at marriage observed for the sample of West African countries. The oldest average age at marriage for males of all 30 countries in this sample is observed for Swaziland at 31.4 years, while the youngest average age at marriage for males in this sample is 21.5 years and is observed for Nepal. Relative to the other South Asian countries in this sample, men in Nepal marry on average 6.4 years earlier than men in India, the country in this regional sample with the eldest average age at first marriage for men. The regional sample with the least within-region variation as measured by the inter-quartile range is South America. Of the ten countries in this sample, there is less than two years difference between the youngest and eldest average age at marriage for men: 24.3 years in Ecuador and 26.0 years in Guyana.

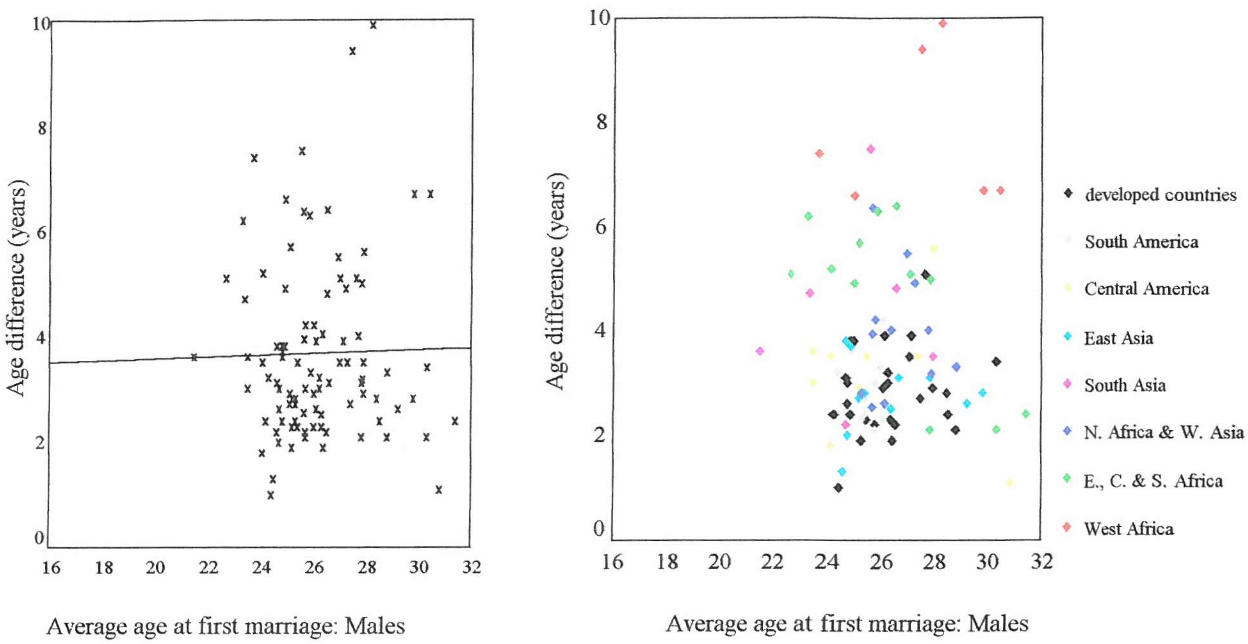
3.4 Association between the age difference and age at marriage at the aggregate level

As discussed in Chapter 1, it has been established that there is an association between the age difference and age at marriage (e.g. Bogue, 1969; Dixon, 1971; Nawar, 1988; Bozon, 1991; Ní Bhrolcháin, 1992). At the aggregate level, there is a greater association with the female age at marriage than with the male age at marriage. This is evident in the Figure 3.3, which also shows the least-squares linear regression line that best fits the data points on the scatterplot.

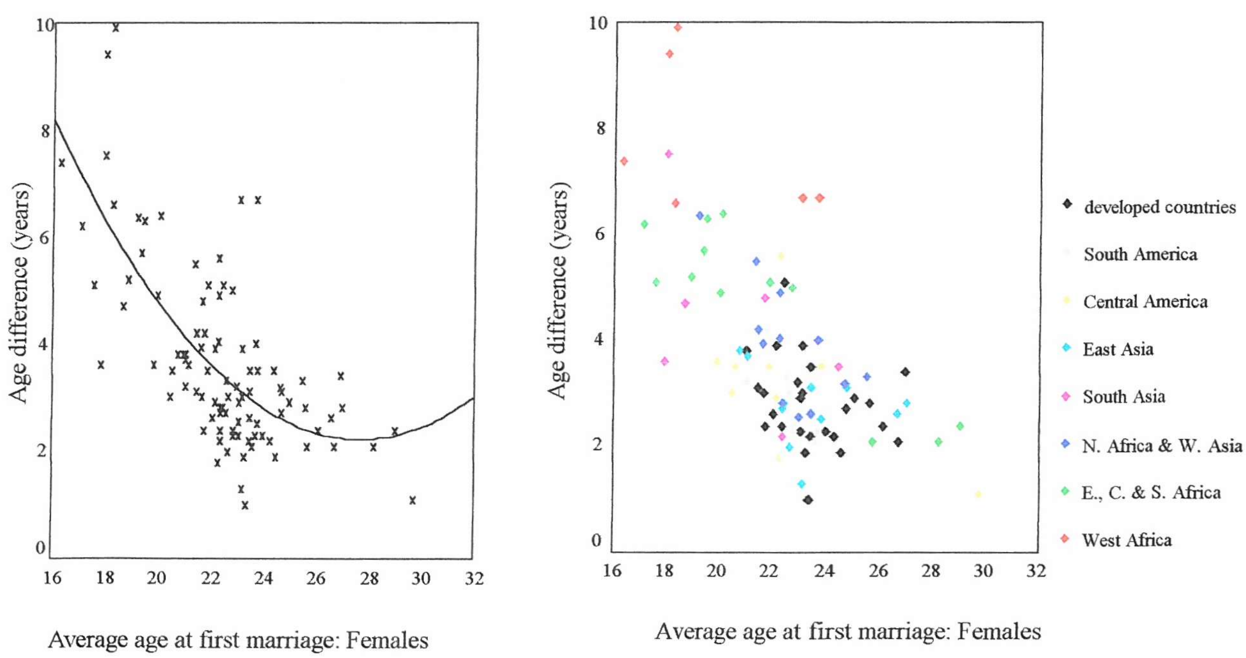
With the average age at first marriage for males predicting the average age difference, the linear regression line is virtually horizontal and the adjusted R-square is .01. Thus, approximately one percent of the variation in the average age difference can be accounted for by the male average age at marriage. When the average age at first marriage of females predicts the average age difference, a slight curvilinear association is identified. Reflecting this in the model involves including a quadratic female age at marriage term in the equation. This results in a statistically significant increase in the adjusted R-square at .49 relative to an adjusted R-square of .43 when the quadratic term is excluded. Figure 3.3 also identifies the countries by UN region for reference.

Figure 3.3 Scatterplots showing the association between the average age difference and the average age at marriage of males and females, respectively

Males



Females



Notes for Figure 3.3

1. E., C. & S. Africa denotes 'East, Central and South Africa'
2. N. Africa & W. Asia denotes 'North Africa and West Asia'
3. See Appendix 2A for list of countries in each region

There is a greater association between the average age difference and the female average age at first marriage rather than the male average age at first marriage, but as Figure 3.3. shows, this association is far from perfect. For example, there are seven countries in this sample where the average age difference is between 6.2 years and 6.7 years, a range of 0.5 years. However, the female age at marriage for these countries ranges from 17.1 years in Ethiopia to 23.7 years in Senegal, a range of 6.6 years. Conversely, there are six countries where female age at marriage is estimated to be between 23.5 years and 24.0 years, a range of 0.5 years, yet the average age difference for these countries lies between 2.1 years in Chile and 6.7 years in Senegal, a range of 4.6 years.

Having identified the main patterns in the regional distribution of the average age difference, the average age at first marriage, and the association between the age difference and its constituent parts, the rest of the chapter examines the hypothesis that the age difference reflects women's status at the aggregate level. Thus, according to the analysis so far, this hypothesis suggests that women's status is relatively low where the average age difference is relatively large, such as in the West African countries studied and also Bangladesh, and that women's status is relatively high where the average age difference is relatively small such as in the developed countries, the East Asian countries, and the Central and South American countries such as Jamaica and Puerto Rico. The next section briefly examines the regional variation in the hypothesised indicators of women's status used to examine this hypothesis. To recap, these are school enrolment at ages 12-17 years, adult illiteracy, and life expectancy at birth, and these are expressed firstly specifically in relation to females and then as the gender difference in the estimates for these indicators.

3.5 Regional variations in the hypothesised indicators of women's status

3.5.1 Indicators specifically in relation to females

Table 3.3 and Figure 3.4 show variation between and within regions for all of the women's status indicators when expressed as indicators specifically in relation to females. Women's status as measured by these indicators is lowest, on average, among the West African countries for all three indicators. Relative to the other regional samples, there is little variation within this region. The median for the two education indicators is lower for the South Asian countries than among the remaining sub-Saharan African countries in the East, Central and South Africa region. However, there is no statistically significant regional variation in female school enrolment or female adult illiteracy between the East, Central and South African countries, the North African and West Asian countries or the South Asian countries (p>.05 according to t-tests). Also similar for these three regions is the extent of variation in these indicators both in terms of the range and the inter-quartile range. The largest range and IQR are both observed for the East, Central and South African countries for female adult illiteracy at 77.8% and 54.8% respectively. Thus, for half of these countries, female adult illiteracy lies between 27.7% and 82.5%. In contrast, there is much less within region variation in female life expectancy. The largest IQR is again observed for the East, Central and South Africa countries at 13.7 years, although the range spans life expectancies found in West Africa (e.g. 46.6 years in Ethiopia and Somalia) to those found in developed countries (e.g. 77.0 years in Reunion).

Women's status as measured by the indicators specifically in relation to females is highest, on average, for the samples of Central and South American, East Asian and developed countries. There is generally little within-region variation for the samples of Central and South American and developed countries. In contrast, while the median for the East Asian sample tends to be on a par with the Latin American and developed country samples, there is much more variation among the sample of East Asian countries for all three indicators.

Of the total variation in the distributions of these female-specific indicators of women's status, most of the variation occurs between regions rather than within-regions: for female school enrolment, 59.3%; for female adult illiteracy, 64.7%; and for female life expectancy, 72.7%.

Table 3.3 Summary statistics for variations in the distributions of three hypothesised indicators of women's status expressed specifically in relation to females, by UN region

Hypothesised indicator women's status by UN Region	Median	IQR	Minimum	Maximum	n
Female school enrolment, ages 12-17 (%)					
West Africa	18.3	7.1	10.0	22.6	6
East, Central and South Africa	39.3	45.8	11.4	79.5	12
North Africa and West Asia	56.5	26.7	30.4	92.9	12
South Asia	22.5	28.0	10.1	64.3	6
East Asia	66.7	45.5	17.4	87.8	11
Central America	64.1	25.8	37.8	80.4	9
South America	70.6	16.8	44.6	91.1	10
Developed countries	85.7	14.8	33.1	100.0	27
Female adult illiteracy (%)					
West Africa	83.2	14.8	74	95	6
East, Central and South Africa	57.9	54.8	19.5	97.3	12
North Africa and West Asia	45.4	36.3	11.3	87.8	12
South Asia	69.5	61.1	16	90.8	6
East Asia	24.7	27.2	6.5	82.0	11
Central America	11.5	27.7	1.4	67.5	9
South America	14.0	12.7	4.5	26.1	10
Developed countries	0.00	4.5	0	35.7	27
Female life expectancy at birth (years)					
West Africa	46.6	4.8	42.6	48.3	6
East, Central and South Africa	53.9	13.7	46.6	77.0	12
North Africa and West Asia	68.5	7.6	60.3	77.4	12
South Asia	56.7	11.2	50.2	72.5	6
East Asia	70.9	7.7	54.7	79.8	11
Central America	74.1	9.1	56.4	77.7	9
South America	70.1	7.0	63.4	75.3	10
Developed countries	78.5	4.7	68.0	81.2	27

3.5.2 The gender difference in indicators

The gender difference for these indicators is calculated as the female estimate minus the male estimate. Table 3.4 shows that in general, the gender difference is more to the female advantage, i.e. narrower for the two education indicators but larger with respect to life expectancy, for countries in the Central and South American samples and the developed countries, than for countries in the other regional samples. There is no statistically significant variation in the gender difference between West Africa, Central, East, and South Africa, North Africa and West Asia, and South Asia for any of the three gender difference indicators ($p<.05$ according to t-tests). The median gender difference in school enrolment is around -12%, i.e.

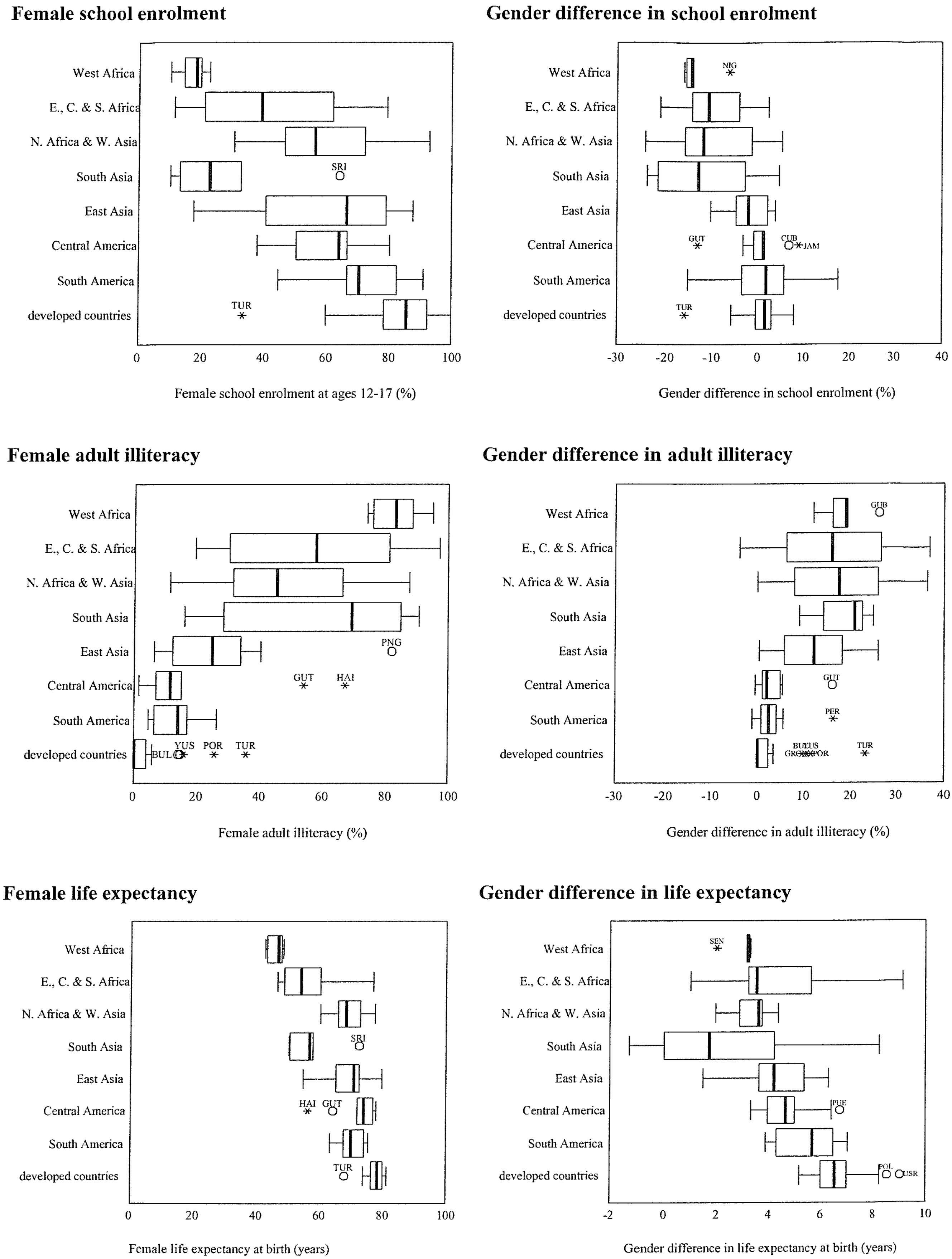
approximately 12% fewer females than males are enrolled in schools, on average in these regions. While the median gender difference in adult illiteracy is around 20% for these four regional samples, i.e. female adult illiteracy is approximately 20 percentage points higher than male adult illiteracy, on average, in these regions. However, there is more variation in the regional medians for life expectancy. While West Africa, Central, East, and South Africa, and North Africa and West Asia, are similar in that female life expectancy at birth is about 3.5 years greater on average, the median for the South Asian sample of countries is smaller at 1.7 years. However, there is much within-region variation especially in the South Asian and Central, East, and South African samples. In contrast, there is little variation in the gender difference in life expectancy for the West African countries, although this is not surprising given that these estimates are from regional model life tables.

Unlike the indicators referring just to females, less than half of the total variation in the distributions of the gender difference indicators occurs between regions, for school enrolment, 41.5%; for adult illiteracy, 45.1%; and for life expectancy, 48.7%. Figure 3.4 shows that there is less variation between regions for the gender difference indicators relative to the indicators specifically in relation to females (see page xv for country codes).

Table 3.4 Summary statistics for variations in the distributions of three hypothesised indicators of women's status expressed as the gender difference, by UN region

Gender difference in: by UN Region	Median	IQR	Minimum	Maximum	n
School enrolment, ages 12-17 (%)					
West Africa	-14.7	3.7	-16.2	-6.4	6
East, Central and South Africa	-10.9	11.6	-21.1	2.3	12
North Africa and West Asia	-12.1	15.2	-24.2	5.4	12
South Asia	-13.1	21.1	-23.7	4.7	6
East Asia	-2.2	7.3	-10.3	3.9	11
Central America	1.2	6.4	-13.2	9.0	9
South America	1.8	9.7	-15.3	17.7	10
Developed countries	1.5	3.9	-15.9	8.0	27
Adult illiteracy (%)					
West Africa	19.0	6.0	12.0	26.2	6
East, Central and South Africa	16.0	21.4	-4.0	36.9	12
North Africa and West Asia	17.4	19.0	0	36.6	12
South Asia	20.9	10.3	9.0	25.0	6
East Asia	12.0	13.3	0.3	26.0	11
Central America	2.0	4.6	-0.4	16.0	9
South America	2.4	3.7	-1.1	16.2	10
Developed countries	0.0	2.5	-0.1	23.3	27
Life expectancy at birth (years)					
West Africa	3.2	0.4	2.0	3.3	6
East, Central and South Africa	3.5	2.6	1.0	9.2	12
North Africa and West Asia	3.6	0.9	1.9	4.4	12
South Asia	1.7	5.5	-1.3	8.2	6
East Asia	4.2	2.0	1.5	6.3	11
Central America	4.6	1.9	3.3	6.7	9
South America	5.7	2.3	3.9	7.0	10
Developed countries	6.5	1.1	5.2	9.0	27

Figure 3.4 Boxplots showing variation in the distributions of three indicators hypothesised as associated with women's status when expressed with respect to females only and also as the gender difference, by UN region



3.5.3 Multidimensional indicators of women's status

As Chapter 2 discussed, some studies have suggested that a single indicator such as female adult illiteracy is unlikely to capture women's status by itself, rather women's status is a multidimensional concept (e.g. Cain, 1984; Caldwell and Caldwell, 1993; Kishor, 1995; Balk, 1997). Thus, in addition to being expressed as the three indicators specifically in relation to females and then as the gender difference for these indicators, women's status is also considered multidimensionally by using a factor analysis of twelve hypothesised indicators of women's status. These hypothesised indicators are shown in Table 3.5.

The Pearson product moment correlation coefficient matrix is examined to ensure correlation between the twelve indicators (see Appendix 3A). All but seven correlation coefficients are greater than ± 0.3. This is confirmed with Bartlett's Test of Sphericity, which examines the hypothesis that the correlation matrix is an identity matrix. The test statistic is large and the corresponding significance level is very small, which indicates that it is unlikely that the population correlation matrix is an identity matrix.

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is also obtained. This is an index for comparing the magnitudes of the observed correlation coefficients to the magnitudes of the partial correlation coefficients. Values for the KMO statistic range from zero to one, with one being the ideal. The KMO in this instance is 0.881, indicating that the correlation between pairs of variables can be accounted for by the other variables.

Three factors are identified from the factor analysis (see section 2.2.5 for justification criteria). The statistically significant factor loadings are given in Table 3.5 ($p<.05$). These are the correlations between each factor and each of the twelve indicators.

Table 3.5 Statistically significant[1] factor loadings obtained from a factor analysis of twelve hypothesised indicators of women's status for 93 countries

Hypothesised indicator of women's status	Factor and loading[2]		
	1	2	3
Female life expectancy at birth	.977		
Maternal mortality ratio	-.928		
% of births attended by a trained birth attendant	.917		
Female enrolment in secondary school	.861		
Female adult illiteracy	-.849		
% of women using contraception	.823		
Total fertility rate	-.776		.303
% of ministerial posts held by women		.952	
Proportion of lower assembly seats occupied by women		.886	
Female economic activity			-.884
Proportion of the population who are Muslim			.767
% of household heads who are female		.484	-.521
% of total variation accounted for by each factor:	**58.7**	**15.3**	**8.3**

Note for Table 3.5:

1. Statistical significance is considered at $p<.05$
2. Factor loadings greater than ±0.3 are considered significant (Child, 1970). See section 2.2.5.

As anticipated from using an oblique method of rotation, there is some correlation between the three factors. This is shown in Table 3.6.

Table 3.6 Correlation matrix for the three factors identified from a factor analysis of twelve hypothesised indicators of women's status for 93 countries

Factor	Factor 1	Factor 2	Factor 3
1	-	-	-
2	.36	-	-
3	-.38	-.40	-

The three factors identified from the factor analysis can be loosely described as:

Factor 1: Demographic and educational benefits to women from social and economic development

The first factor accounts for over half of the total variance and is composed of seven variables. Four variables have a significant positive association with this factor: female life expectancy, contraceptive use, the presence of birth attendants, and female secondary school enrolment. Three variables have a significant negative association with this factor: total fertility rate, maternal mortality ratio, and female adult illiteracy. Thus, an increasingly positive score for this factor denotes higher women's status according to this measure. In the text and plots this factor is referred to as 'women's status and development'.

Factor 2: Women's status in the public domain

The second factor accounts for about 15% of the total variance and is composed of three variables. All three of the variables have a significant positive association with this factor: Representation of women in parliamentary assembly, representation of women in ministerial posts, and female household headship. Thus, an increasingly positive score for this factor denotes higher women's status according to this measure. In the text and plots this factor is referred to as 'women's status in the public domain'.

Factor 3: Women's status in terms of women's traditional roles

The third factor accounts for less than 10% of the total variance and is composed of four variables. Two variables have a significant positive association with this factor: total fertility rate, and the proportion of the population who are Muslim. Two variables have a significant negative association with this factor: female economic activity, and female household headship. Thus, unlike factors one and two, an increasingly *negative* score for factor three denotes higher women's status according to this measure. In the text and plots this factor is referred to as 'women's status and traditionalism.'

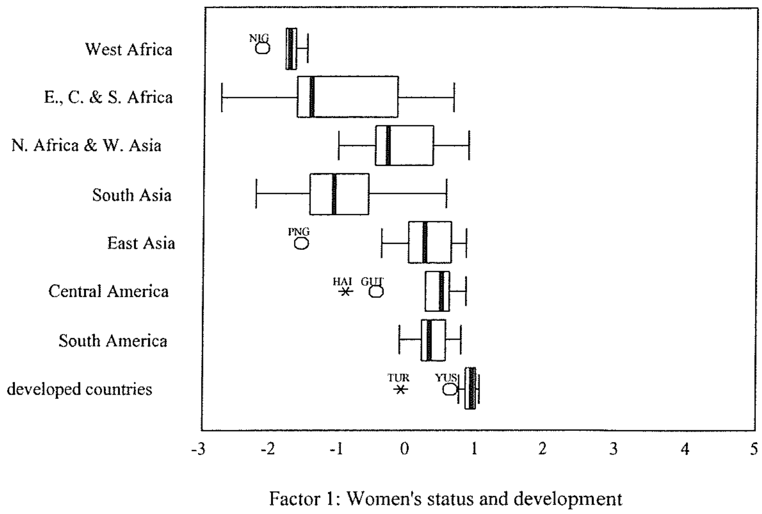
Table 3.7 gives summary statistics for variations in the distributions of the three women's status factors by UN region and Figure 3.5 displays these variations graphically (see page xv for country codes).

Table 3.7 Summary statistics for variations in the distributions of three women's status factors by UN region

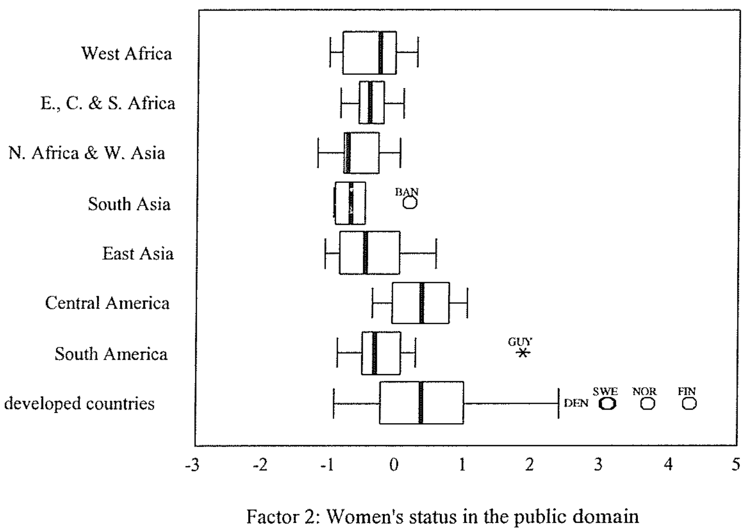
Women's status factor: by UN region	Median	IQR	Minimum	Maximum	n
'Women's status and development'					
West Africa	-1.76	0.28	-2.18	-1.49	6
East, Central and South Africa	-1.43	1.68	-2.75	0.64	12
North Africa and West Asia	-0.29	0.89	-1.03	0.87	12
South Asia	-1.10	1.33	-2.24	0.54	6
East Asia	0.24	0.85	-1.57	0.84	11
Central America	0.48	0.79	-0.90	0.83	9
South America	0.31	0.45	-0.11	0.77	10
Developed countries	0.92	0.15	-0.09	1.04	27
'Women's status in the public domain'					
West Africa	-0.29	0.92	-1.03	0.27	6
East, Central and South Africa	-0.42	0.39	-0.87	0.08	12
North Africa and West Asia	-0.75	0.56	-1.20	0.04	12
South Asia	-0.70	0.61	-0.97	0.19	6
East Asia	-0.48	0.89	-1.07	0.57	11
Central America	0.35	0.87	-0.38	1.05	9
South America	-0.34	0.66	-0.87	1.88	10
Developed countries	0.36	1.40	-0.93	4.29	27
'Women's status and traditionalism'					
West Africa	0.49	1.25	-0.70	1.56	6
East, Central and South Africa	-0.27	1.05	-1.00	0.64	12
North Africa and West Asia	1.92	2.48	-0.09	2.38	12
South Asia	0.41	1.97	-0.65	2.22	6
East Asia	-0.62	1.48	-1.34	1.04	11
Central America	-0.15	0.71	-1.53	0.86	9
South America	0.00	0.42	-0.36	0.49	10
Developed countries	-0.80	0.50	-1.41	1.11	27

Figure 3.5 Boxplots showing variation in the distributions of three women's status factors, by UN region

Women's status and development factor



Women's status in the public domain factor



Women's status and traditionalism factor

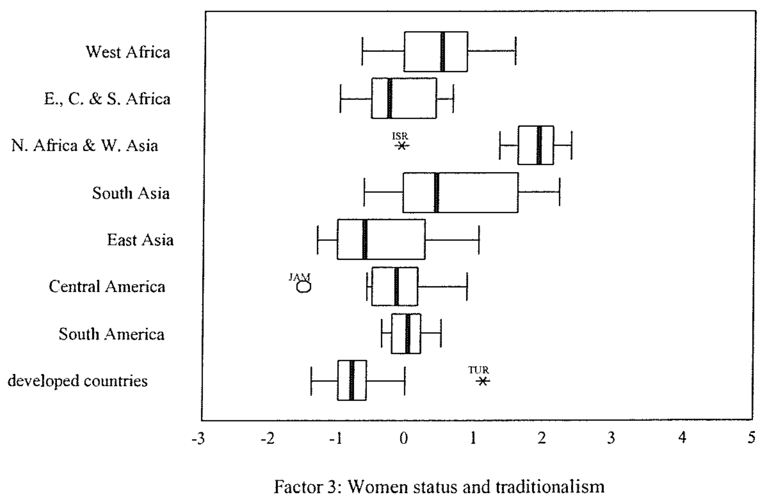


Figure 3.5 shows variation between and within regions for all three women's status factors, although the extent of these variations varies by factor. For factor 1, women's status and development, the West African countries have the lowest median score and thus, the lowest women's status according to this measure. Although statistically, there is no significant difference between the scores observed for the sample of West African countries and the sample of countries in the rest of sub-Saharan Africa or the South Asian countries (p>.05 according to the Mann-Whitney test since variance homogeneity can not be assumed).

There is no statistically significant variation in the scores obtained for factor 2, women's status in the public domain, between the regional samples, apart from the sample of Central American countries (p>.05 according to the Mann-Whitney test since variance homogeneity can not be assumed). Indeed, of the total variation in this factor, less than one-third is due to variation between regions, i.e. more than two-thirds of the observed variation occurs within regions. This compares to factors 1 and 3 where the opposite pattern is observed, i.e. approximately two-thirds of the total variation in these factors is due to variations between regions. One sample where there is much within-region variation in the scores for factor 2 is the sample of developed countries. Four of the developed countries have a score for this factor in excess of 2, which as Figure 3.5 shows, is relatively large. This reflects the significance of the two political variables used to derive this factor, and that in all four Scandinavian countries, women hold at least one-third of lower assembly parliamentary seats and at least 29.7% of ministerial posts held.

According to factor 3, the regional sample with the lowest status, represented in this case by an increasingly positive score, is North Africa and West Asia. This is mainly due to the fact that the proportion of the population who are Muslim has the second largest loading for this factor, and it is estimated that at least 85% of the population in all but one of the countries in this regional sample is Muslim. Indeed, there are only eight other countries in the other seven regions with a similar proportion of the population that is Muslim.

This brief overview of the regional variation in these three factors of women's status shows how different dimensions of women's status can result in different regional patterns, reflecting the multidimensionality of women's status. Although, it is important to recall that these factors are derived from a far from exhaustive selection of hypothesised indicators, such that the inclusion of other indicators may result in different and/or additional factors. The next sections examine the hypothesis that the age difference reflects women's status at the aggregate level by examining the association between the average age difference and the hypothesised indicators of women's status discussed above.

3.6 Examining the hypothesis that the age difference reflects women's status at the aggregate level using indicators specifically in relation to females

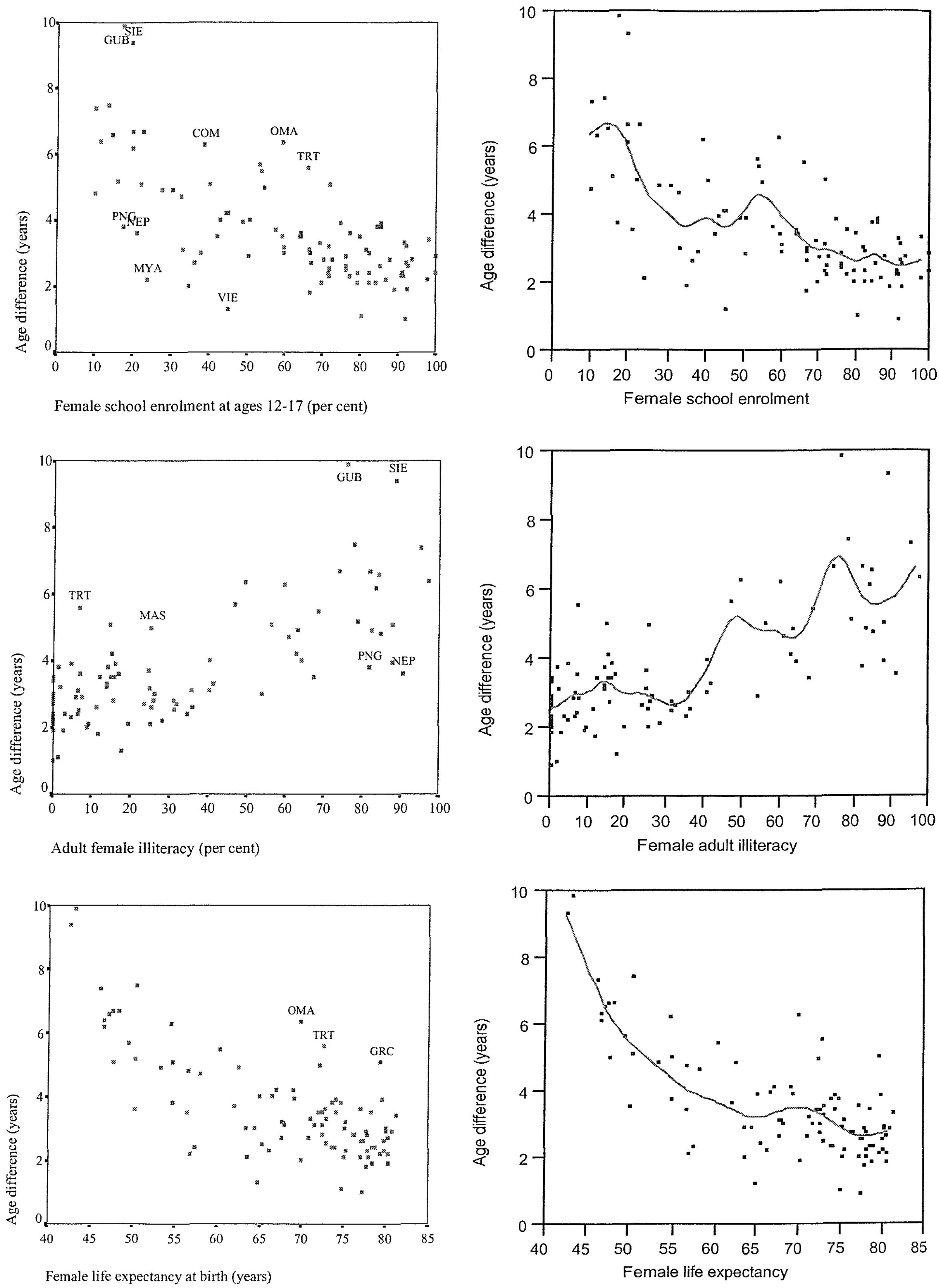
This section investigates the hypothesis that the age difference is an indicator of women's status at the aggregate level by examining scatterplots of the association between the unidimensional indicators described above and the average age difference. First, the association between the average age difference and the women's status indicators that refer only to females is examined. Figure 3.6 shows the corresponding scatterplots. Using the statistical package, JMP (see section 2.2.6), a spline is fitted to each plot with lambda equal to 100, and the outlying countries, which are defined as those countries with a standardised residual from the spline of greater than ±1.96, are identified (see page xv for country codes used for all plots).

3.6.1 Female school enrolment

The spline in Figure 3.6 shows some evidence of a negative association, albeit weak, between female school enrolment and the average age difference among this study's sample of countries. 57.7% of the variation in the average age difference is accounted for by female school enrolment. The unaccounted for variation in the average age difference is particularly evident for countries where less than about half of girls aged 12-17 are enrolled in school. For example, enrolment is of a similar proportion in Myanmar and Guinea-Bissau at 23.9% and 17.1% respectively, yet the average age difference is 2.2 years in Myanmar but almost 10 years in Guinea-Bissau. Where enrolment is greater than 60%, there is less heterogeneity in the average age difference and a plateau forms of countries with average age differences generally between two and four years. Exceptions are Trinidad and Tobago and Greece, where the average age difference is greater than four years, and Puerto Rico, Jamaica and Ireland where the average age difference is less than two years. Despite this variation in the age difference, it is interesting to note that Trinidad and Tobago and Jamaica have similar levels of enrolment with approximately two-thirds of girls aged 12-17 enrolled in school.

Figure 3.6 Scatterplots showing the association between the average age difference and the indicators of women's status specifically in relation to females

With outlying countries identified: **With a smoothing spline fitted (lambda=100):**



3.6.2 Female adult illiteracy

The spline in Figure 3.6 fluctuates in the scatterplot of the average age difference plotted against female adult illiteracy. With lambda equal to 100, female adult illiteracy accounts for 65.3% of the variation in the average age difference. There is much variation in the age difference for all levels of illiteracy, which can be considered as two groups. One group consists of relatively more developed countries, including some Caribbean, East European and East Asian countries, with relatively low female adult illiteracy (less than 40%), and where age differences tend to be smaller, on average (less than five years). Although there are still some exceptions. For example, approximately one in six adult women are estimated to be illiterate in both Greece and Vietnam, yet the average age difference for Greece is more than three times the size of the average age difference in Vietnam, at 5.1 years and 1.3 years respectively.

Sub-Saharan African and South Asian countries with high female illiteracy, 70% or more, tend to have large average age differences. For example, Benin has an average age gap of 6.6 years while 84.4% of women are illiterate. Similarly, the average age gap in Bangladesh is 7.5 years while 78.8% of women are illiterate. However, there is also heterogeneity in the average age difference when illiteracy is high. In this group, there are countries such as the Comoros and Mauritania where both illiteracy and the average age difference are relatively high, but there are also countries such as Guatemala where 54.5% of adult females are illiterate, but the age gap is relatively narrow at three years, on average. Nepal is an even more extreme case. The average age gap is again small at 3.6 years but 90.8% of women are estimated as illiterate.

3.6.3 Female life expectancy

The smoothest of the three splines is observed for the scatterplot of the association between the average age difference and female life expectancy at birth. However, a similar proportion of the variation in the average age difference is accounted for by female life expectancy as is accounted for by female adult illiteracy (64.2%). There is some clustering of relatively more developed countries where female life expectancy is at least 70 years and the average age difference is relatively small (less than four years). However, as Figure 3.6 identifies, there are outlying countries where female life expectancy is high but the average age gap is not as narrow as would be expected for this level of life expectancy: Oman (6.4 years), Trinidad and Tobago (5.6 years), and Greece (5.1 years).

As illustrated by the spline, the gradient of this negative association is greatest where female life expectancy is less than 70 years. The countries where female life expectancy is at its lowest (less than about 50 years) are also the countries where the average age gap is at its widest (more than six years). These are all sub-Saharan African countries, with the one exception of Bangladesh. However, there is much heterogeneity in this part of the plot. For example, the average age difference is of a similar magnitude in Mozambique and Egypt at 5.1 years and 5.5 years respectively, yet there is a 12 year difference in the corresponding life expectancies, 47.8 years and 60.3 years respectively.

Having identified some association, albeit weak, between the average age difference and the three hypothesised indicators of women's status that refer just to females, the next section repeats this analysis but with the indicators expressed as the gender difference.

3.7 Examining the hypothesis that the age difference reflects women's status at the aggregate level using the gender difference in the three indicators

Figure 3.7 shows scatterplots of the association between the average age difference and the three indicators of women's status, expressed as the arithmetic gender difference, which is calculated as the female estimate minus the male estimate. A spline is fitted to each plot with lambda equal to 100, and the outlying countries from this spline are again identified.

3.7.1 Gender difference in school enrolment

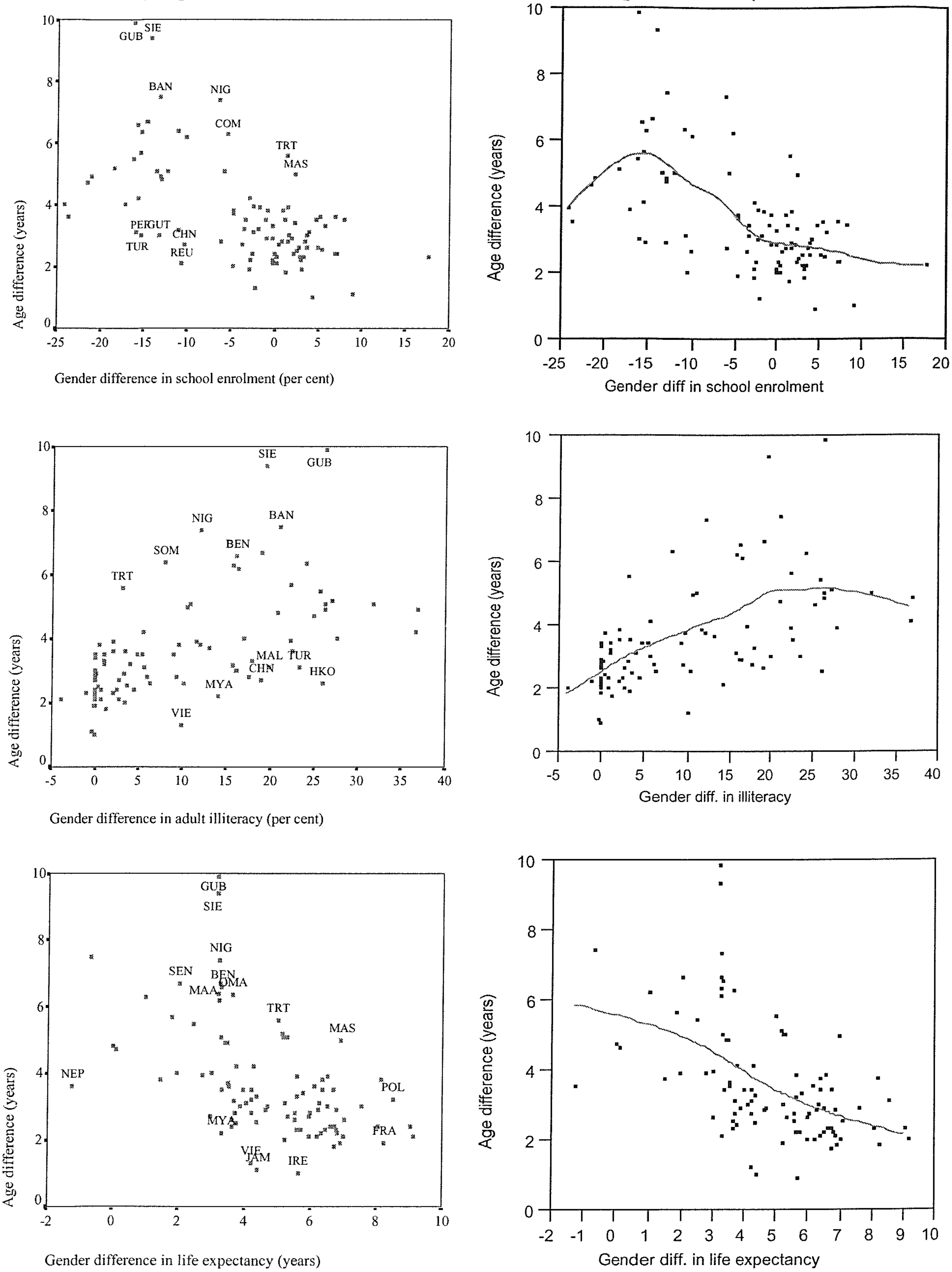
The spline in Figure 3.7 suggests a curvilinear association between the gender difference in school enrolment and the average age difference. The corresponding adjusted R-square for this plot is 0.44. Thus, more of the variation in the average age difference remains unaccounted for by this indicator of women's status than is actually accounted for. There is a cluster of countries where the gender difference is between ±5% and most average age differences are between two and four years. Exceptions are Ireland, Jamaica and Vietnam where average age gaps are one year smaller than this, and Mauritius and Trinidad and Tobago where average age gaps are about one year larger than this. It is interesting to note Uruguay, which has a relatively small average age difference at 2.3 years, and which has the greatest positive gender difference at 17.7%, i.e. the school enrolment rate for females aged

12-17 is 17.7 percentage points greater than the school enrolment rate for males aged 12-17. As the plot shows, there is increasing variation in the average age difference as the gender difference becomes negative, in other words, as the proportion of males aged 12-17 enrolled exceeds the proportion of females aged 12-17 enrolled. The spline suggests that the countries with the greatest negative discrepancy between the proportion of males and females enrolled are not those with the largest average age differences. For example, both Nepal and Iraq have gender differences in excess of –20%, yet have average age gaps which are typical in magnitude of more developed countries at 3.6 and 4.0 years respectively. In comparison, those countries with the largest average age gaps, such as Guinea-Bissau, Sierra Leone, Mauritania, Senegal, and Bangladesh, have gender differences that are between five and ten percentage points smaller, at around -15%.

This observation may be considered as evidence in support of a hypothesis that as a society develops economically, women's position relative to men's position deteriorates before it improves, at the aggregate level (Boserup, 1970). It has been argued that the societal shift from labour-intensive traditional agricultural systems to modern industrialised methods emphasises the importance of education and training for obtaining employment outside of the agricultural sector. However, at the aggregate level, females miss out on educational opportunities due to limited resources, and since women no longer have a role in household production, their roles are limited to those of wives and mothers. Only with time, as the importance of educating females as well as males is incorporated into developmental policy, and more resources become made available, do women begin to catch-up with men, and thus the gender difference in school enrolment and illiteracy narrows. Thus, the average age difference in Figure 3.7 may be acting as a proxy for the level of socio-economic development in a society. This in turn reflects the aggregate-level association between the age difference and female age at marriage, which as Chapter 2 discussed (section 2.2.4), is sometimes hypothesised as a measure of socio-economic development (e.g. UN, 1987). Of course, this is an over-simplification of what is a complex process in which a number of factors are involved. Furthermore, this hypothesis is very tentatively assumed since the conclusion is based on a small number of cases. Indeed, since the sample of 93 countries comes from a population of 212 countries (see Appendix 2A), it is reasonable to question whether there may have been more or less evidence for Boserup's (1970) hypothesis if data had been available for more countries.

Figure 3.7 Scatterplots showing the association between the average age difference and the indicators of women's status expressed as the gender difference

With outlying countries identified: **With a spline fitted (lambda=100):**



3.7.2 Gender difference in adult illiteracy

Applying a spline to the scatterplot of the gender difference in adult illiteracy and the average age difference also reveals a curvilinear association. The corresponding R-square is 0.38, thus approximately two-thirds of the variation in the average age difference is unaccounted for by this indicator of women's status. There is a cluster of countries where the average age gap is small (less than about four years) and the gender difference is small (less than about 10%). This cluster is made up of Central and South American countries (e.g. Mexico, Puerto Rico, Colombia), and East European countries (e.g. Hungary, Poland), as well as some of the Middle Eastern (e.g. Kuwait, United Arab Emirates) and East Asian countries (e.g. the Philippines and Thailand). Many of the Asian countries in this sample have average age differences of between two and four years yet the corresponding gender difference in illiteracy ranges from almost no difference to more than 25%. For example, the gender difference in the Philippines is 0.3% but in Hong Kong it is 26.0%, yet the average age gap for both countries is about 2.5 years. This dispersion in the gender difference also applies to countries with large average age differences. Most sub-Saharan African countries have average age gaps of between five and seven years, yet the gender difference in illiteracy ranges from just over 5% to almost 40%. As observed for the gender difference in school enrolment, some of the countries with the widest gender differences, such as Syria, Zaire and Mozambique, have average age differences that are similar in magnitude to those observed in some developed countries. While the countries with the largest average age differences have gender differences in illiteracy that are similar in magnitude to some of the more developed countries, such as China, Hong Kong, and Turkey where age gaps are less than three years, on average. This may also be considered as evidence in support of Boserup's (1970) hypothesis.

3.7.3 Gender difference in life expectancy

The spline in Figure 3.7 suggests an almost linear association for the scatterplot of the gender difference in life expectancy and the average age difference. However, only 29.8% of the variation in the average age difference can be accounted for by the gender difference in life expectancy. It is possible to consider this distribution in terms of two groups of countries. Where the gender difference is less than about 3.5 years, there is much heterogeneity in the average age difference. For example, at one extreme is Nepal, where the average age difference is smaller than would be expected (3.6 years) given the negative gender difference (i.e. male life expectancy is greater than female life expectancy). While at the other extreme is a band of sub-Saharan African countries where age differences tend to be larger, on average, but the gender difference is similar to that observed for some Caribbean countries (e.g. Haiti and Cuba), and some West Asian countries (e.g. Syria and Iraq). In countries where the average age difference is relatively narrow, less than about four years on average, there is much variation in the gender difference although there are some regional patterns. Many of the Asian countries in this sample cluster where female life expectancy is between about 3.5 and five years greater than male life expectancy (e.g. Sri Lanka and the United Arab Emirates). In contrast, there are mainly more developed countries where the gender difference is in excess of five years (e.g. Ireland and Poland), while Central and South American countries tend to be spread across this range in the gender difference (e.g. Colombia and Guatemala).

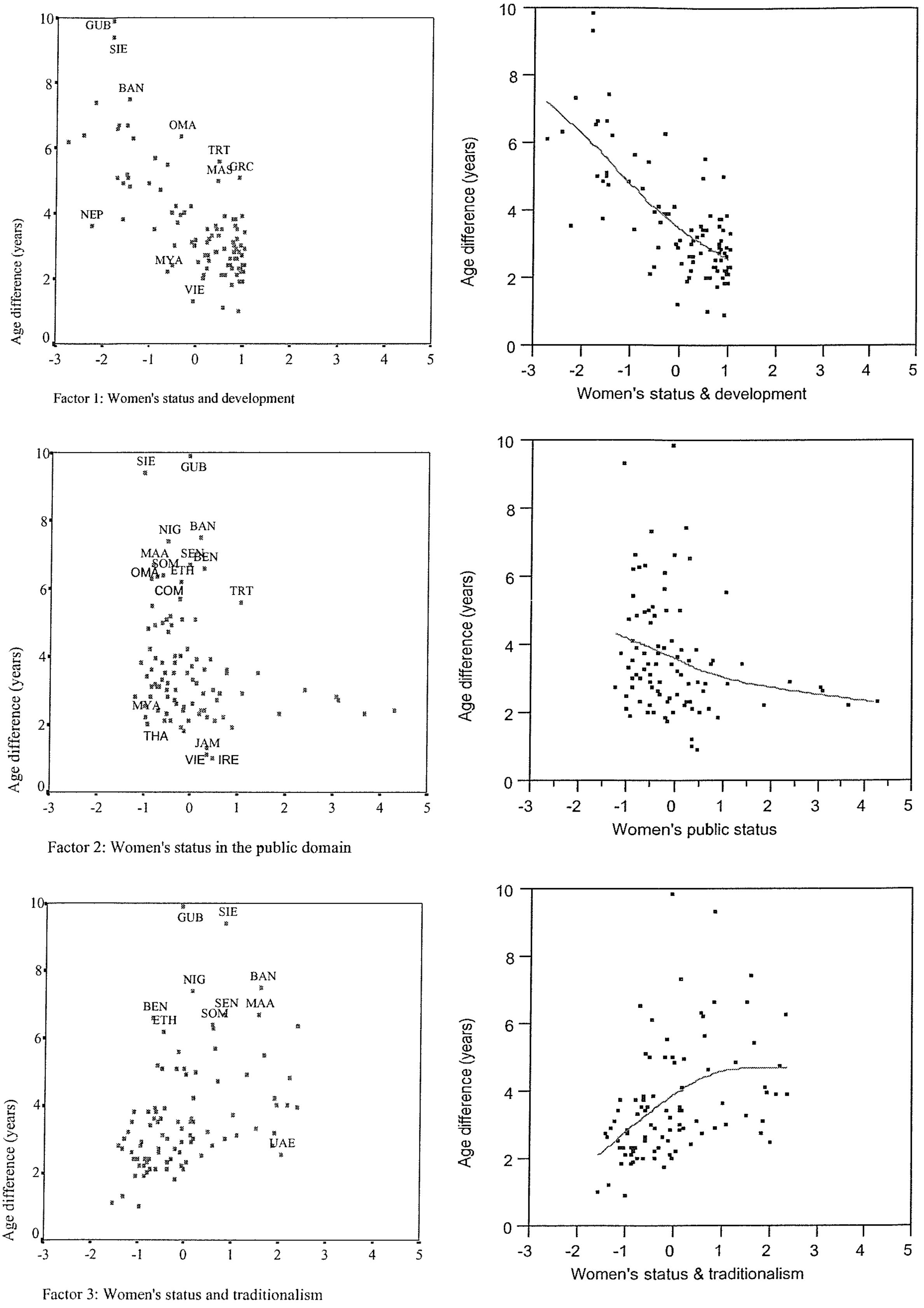
Having observed evidence of an association that is in the hypothesised direction between the average age difference and the unidimensional indicators of women's status, the next section investigates whether such an association exists between the average age difference and the three multidimensional factors of women's status.

3.8 Examining the hypothesis that the age difference is an indicator of women's status at the aggregate level using multidimensional factors relating to women's status

Figure 3.8 shows the association between the average age difference and each of the three women's status factors described in section 3.5.3. A spline is fitted to each plot with lambda equal to 10, and the outlying countries from this spline are again identified (see page xv for country codes).

Figure 3.8 Scatterplots showing the association between the average age difference and factors obtained from a factor analysis of 12 hypothesised indicators of women's status

With outlying countries identified: **With a spline fitted (lambda=10):**



Of the three plots presented in Figure 3.8, the greatest association seems to be between the average age difference and factor 1, 'women's status and development', although, there is still much variation in the average age difference for all values of this factor. From fitting smoothing splines with lambda equal to ten, over half of the variation in the average age difference can be accounted for by this factor (R-square equals 0.56). This compares to the plot of factor 2 for which it is difficult to identify any association (R-square equals 0.07), and for factor 3, for which a slight curvilinear association is evident (R-square equals 0.21).

For factor 1, the association is in the hypothesised negative direction. Thus, as the women's status and development factor tends towards a positive score, the average age difference narrows. Apart from Oman, all of the countries in this sample with wide average age differences (greater than six years) score less than −1 for factor 1. This suggests that in these countries women's status is low according to this factor. However, there are several countries in this sample where women's status is equally low according to this factor, but where average age gaps are generally a couple of years smaller. These countries include Nepal (identified on the plot as an outlier), Papua New Guinea, Mozambique, C.A.R., Zaire, Congo, Pakistan, and Morocco. The opposite pattern is also evident, albeit to a lesser extent. For example, in Greece, Mauritius, and Trinidad and Tobago, women's status according to this measure is comparable to women's status in many of the developed countries in this sample, yet the average age difference is larger in the former countries at around five years. Indeed, if a spline is fitted to just positive factor scores, then it is horizontal with a corresponding R-square value of 0.009, confirming a lack of association with this indicator of women's status when the factor scores are constrained in such a way that the average age differences are typically small.

Evidence of a very slight association in the hypothesised negative direction can be identified between factor 2, women's status in the public domain, and the average age difference. Although, the majority of countries in this sample have a factor score of between -1 and +1 with average age differences ranging from approximately two to eight years. The only exceptions are Guyana and five of the north-west European countries in this sample: the Netherlands, Denmark, Sweden, Norway and Finland, which score very highly in terms of women's status in the public domain and which have narrow average age differences. The almost flat spline and the corresponding value for R-square, 0.07, both reflect this lack of association.

The spline suggests a very slight positive curvilinear association between the average age difference and factor 3, 'women's status and traditionalism'. Thus, in most of the countries where the average age difference is less than four years, women's roles are relatively less traditional, which is represented in this plot by a negative factor score. Conversely, in most of the countries where the average age difference is greater than six years, women's roles are relatively more traditional, represented by a positive factor score. As with factor 1, there is much variation in this pattern. For example, according to this measure, women's status in Benin, Ethiopia, C.A.R., the Congo and Mozambique, is of a comparable level of traditionalism to some developed countries such as Italy, Japan, Sweden and the Netherlands, yet the average age difference tends to be three or four years smaller in these developed countries. At the other end of this factor's scale are a number of West Asian countries such as the United Arab Emirates, Kuwait, and Algeria where average age differences tend to be less than four years, but where factor scores are large, which suggests a greater social expectation for women to become wives and mothers. Thus, if these Islamic countries and the West African countries are ignored, the association between the average age difference and this indicator of women's status is slightly more evident.

3.9 Examining the aggregate-level association between the age difference and hypothesised indicators of women's status, net of female age at marriage

Having identified evidence of an association between the average age difference and each of the hypothesised uni- and multi-dimensional indicators of women's status, this section attempts to control for the association between the age difference and female age at marriage observed in Figure 3.3 by using the 'net age difference'. The net age difference is quantified as the standardised residuals from regressing the male age at marriage on the female age at marriage (and the female age at marriage squared). As shown in Chapter 2 (section 2.2.6), this is a simplification of the equation where the age difference is the dependent variable and the female age at marriage and the female age at marriage squared are the explanatory variables. The standardised residuals range from –2.32 for Nepal to +3.16 for Guinea-Bissau. Thus, the analysis based on female age at marriage over-estimates the male age at marriage and thus the age difference for Nepal, while for Guinea-Bissau, the analysis based on female age at marriage under-estimates the male age at marriage and thus the age difference. Scatterplots of the association between the net age difference and each of the women's status indicators are now presented (see page xv for country codes).

Figure 3.9 Scatterplots showing the association between the 'net age difference' quantified as standardised residuals and the hypothesised indicators of women's status specifically in relation to females

With outlying countries identified: **With a spline fitted (lambda=100):**

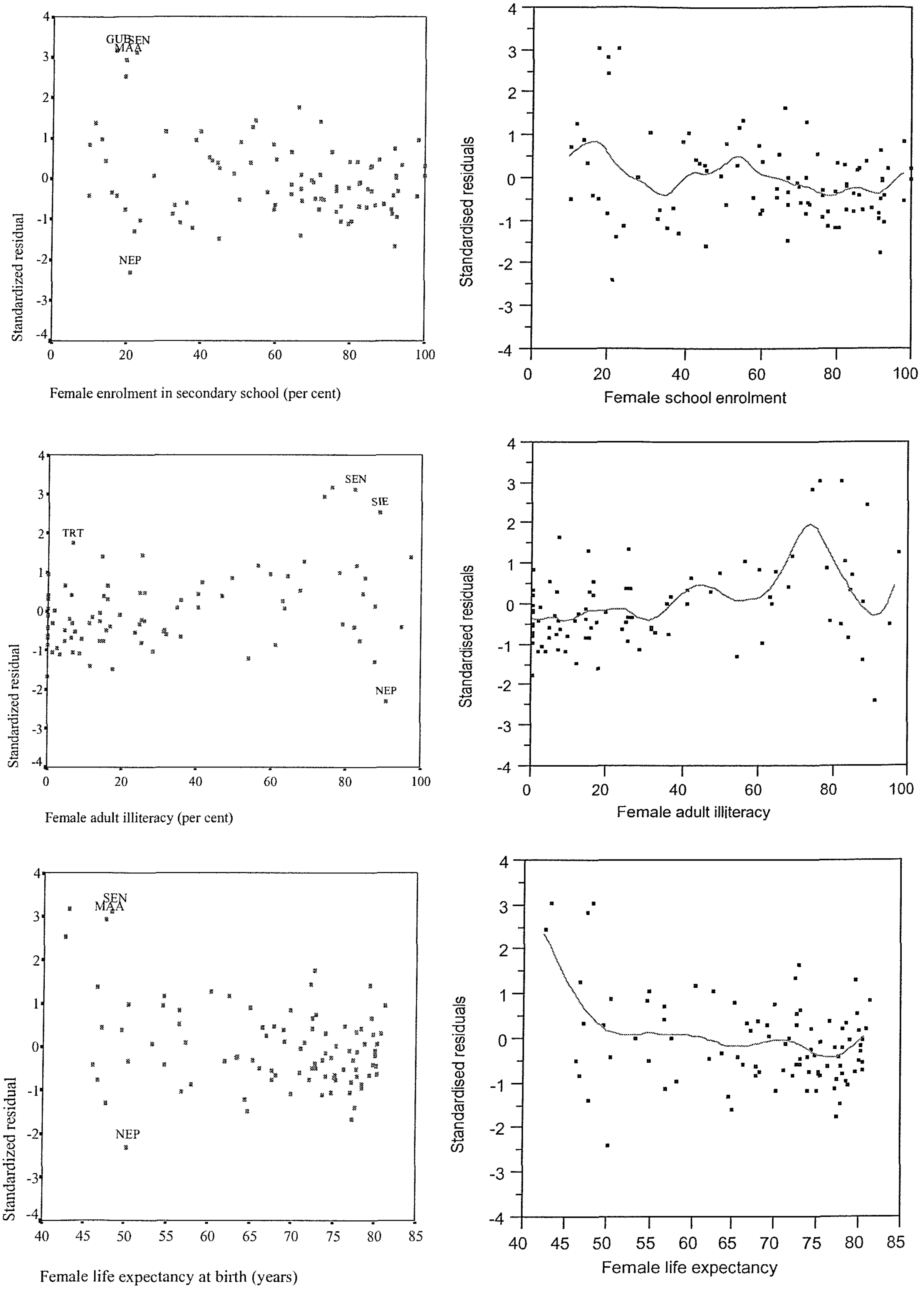


Figure 3.10 Scatterplots showing the association between the 'net age difference' quantified as standardised residuals and the hypothesised indicators of women's status expressed as the gender difference in these indicators

With outlying countries identified: **With a spline fitted (lambda=100):**

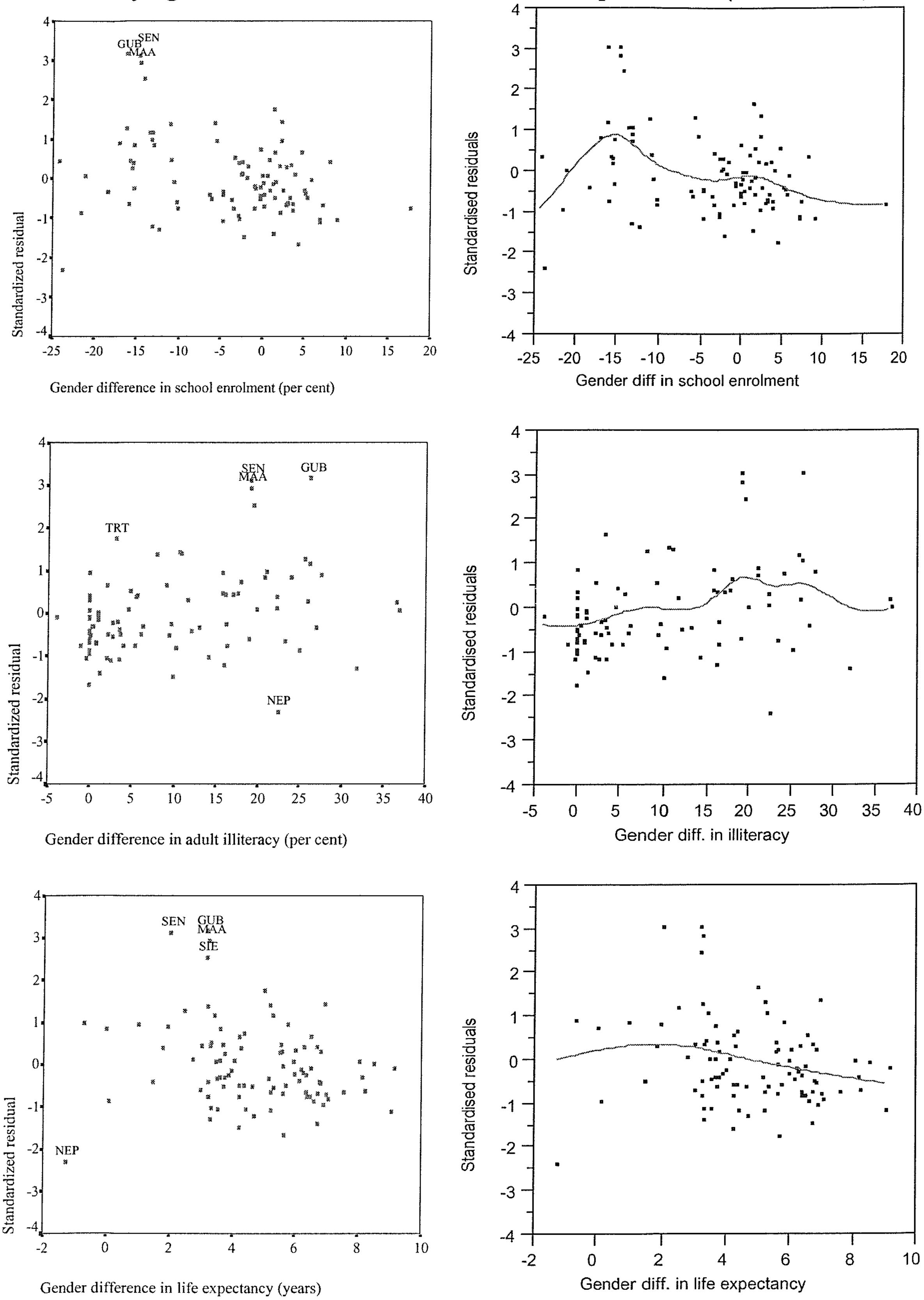
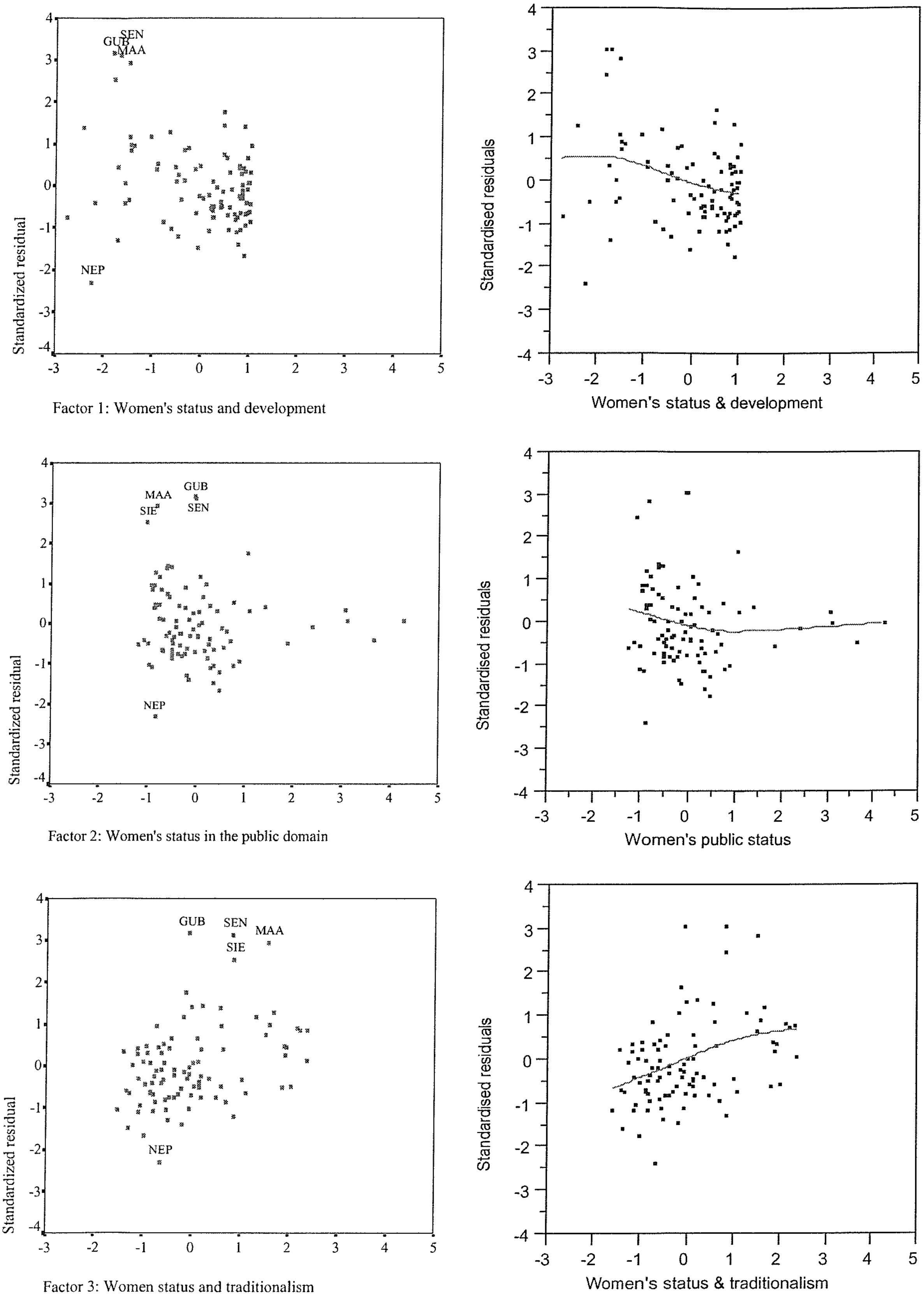


Figure 3.11 Scatterplots showing the association between the 'net age difference' quantified as standardised residuals and factors obtained from a factor analysis of twelve hypothesised indicators of women's status

With outlying countries identified: **With a spline fitted (lambda=10):**



The plots in Figures 3.9, 3.10, and 3.11 show the association between the various indicators of women's status and the age difference net of female age at marriage, expressed as the standardised residuals from regressing the average male age at marriage on the average female age at marriage and female age at marriage squared. In most of the plots, it is difficult to identify any particular association. Fitting a smoothing spline results in a fluctuating line, especially in the plots for the indicators that refer just to females (Figure 3.9). The splines fluctuate much less in the plots of the gender difference in indicators (Figure 3.10), and even less in the plots of the factors of women's status (Figure 3.11). Although, there is evidence of a slight negative slope as the gender difference becomes increasingly to the males' advantage in the plot of the gender difference in school enrolment, which suggests that the net age difference is smaller where women's status is lowest, as measured by the gender difference. However, in light of the small number of countries that this observation is based upon, this is a very tentative observation. As the plots suggest, the values for the adjusted R-squared statistic from regressing the standardised residuals on the women's status indicators are seldom large. They range from 0.00, thus none of the variation is accounted for when the 'women's public status' factor is used as the explanatory variable, to 0.13, thus approximately an eighth of the variation is accounted for when the 'women status and traditionalism' factor is used as the explanatory variable.

If the five countries identified in the plots as outliers (Nepal, Sierra Leone, Mauritania, Guinea Bissau, and Senegal) are excluded from the regression of the male age at marriage on the female age at marriage (and the female age at marriage squared), and new standardised residuals are obtained then the fluctuations in the spline are much reduced in all of the plots of the association with the various indicators of women's status (not shown). Thus, the fluctuations in the splines observed in Figures 3.9, 3.10 and 3.11 may be partly due to the influence of a handful of, mainly, West African countries. Although, it is worth noting that while these countries are outliers in terms of the average age difference, they are seldom outliers in terms of the indicators of women's status. Furthermore, these countries are from incomplete regional samples. Thus, if this analysis was repeated with complete samples then these countries may not appear as outliers.

3.10 Discussion and conclusions

There were two main objectives of the analysis presented in this chapter. First, this chapter sought to examine regional variations in the average age difference and the average age at first marriage and also to consider the association between these variables. The average age at first marriage for both males and females was estimated with the singulate mean age at marriage (SMAM), such that the average age difference at first marriage was calculated as the male SMAM minus the female SMAM. In general, a slightly larger proportion of the total variation in the average age difference occurred between regions rather than within regions. It seems that this result was due to the statistically significantly larger average age differences that were observed for the sample of West African countries. Bangladesh was identified as an interesting country since it is the only country in this sample not in sub-Saharan Africa with an average age difference as large as observed for the West African countries.

The second objective of this chapter was to examine the hypothesis that the average age difference reflects women's status at the aggregate level. In this chapter, this hypothesis was examined by considering the association between the average age difference and hypothesised indicators of women's status from the UN's WISTAT3 database: school enrolment at ages 12 to 17, adult illiteracy, and life expectancy at birth. These indicators were chosen because they were considered to strike a balance between comprehensive coverage and cultural generalisability. Even so, it may be argued that the level of adult illiteracy is not an appropriate indicator of women's status in more developed countries where it is negligible. However, relating this chapter's selection of indicators to Mason's (1986) list of commonly used indicators (Table 2.2), the indicators used for this chapter's aggregate-level analysis are clearly some of the more appropriate and more readily available indicators. In addition to expressing the three indicators specifically in relation to females, and then as the gender difference for these indicators, women's status was also considered multidimensionally with three factors that were identified from a factor analysis of twelve hypothesised indicators of women's status. This is because some studies have suggested that it is unlikely that a single indicator such as female adult illiteracy can capture women's status by itself, rather women's status is a multidimensional concept (e.g. Cain, 1984; Caldwell and Caldwell, 1993; Kishor, 1995; Balk, 1997).

Regardless of the hypothesised indicator of women's status used or how it was expressed, only a weak association was observed with the average age difference. Some evidence of a slight curvilinear association was identified in Figure 3.7 between the average age difference and the gender difference in the educational indicators. This was considered as evidence in support of Boserup's (1970) hypothesis that, as a society develops, the status of women declines before it improves. However, this observation was tentatively suggested given the incomplete samples and the relatively small number of countries upon which the observation was based.

This chapter's analysis concluded by attempting to quantify the 'net age difference' to try to take into account the associations between (1) the average age difference and the average age at marriage of females (e.g. Bogue, 1969), and (2) the average age at marriage of females and the indicators of women's status (e.g. Mason, 1985, 1986). This was estimated by using the residuals from regressing the male SMAM on the female SMAM and the female SMAM squared since as discussed in Chapter 2, this is more statistically robust than regressing the age difference on the female SMAM and the female SMAM squared (Cain *et al*, 1992; Yanez *et al*, 1998), as well as being conceptually simpler (Plewis, 1985). From this analysis, little association was evident between the average age difference and the hypothesised indicators of women's status once the age difference was expressed net of the female average age at marriage. Thus, at least in a statistical sense, it seems that the aggregate-level association between hypothesised indicators of women's status and the age difference is spurious due to (1) the association between the average age at marriage of females and the average age difference, and (2) the association between female age at marriage and hypothesised indicators of women's status.

The extent to which these associations can be teased apart in any causal sense was not considered. Thus it is not possible to deduce from this analysis that women's status is in some sense prior to age at marriage or the age difference, or vice versa. This is partly because of the aggregate, cross-sectional nature of the data used for this analysis, which are not suitable for identifying the processes that operate between women's status, female age at marriage, and the age difference, such as the direction of wealth transfers at marriage or the marriage arrangers. As Obermeyer (1992, p. 38) described, aggregate-level analyses are of *"limited usefulness when attempting to understand demographic change."* However, what can be concluded from these results is that relative to the average age at marriage for females, the average age difference does not add anything in a statistical sense when considering the association with women's status in aggregate-level analysis. If this is the case then cross-national studies of women's status may be liberated in terms of their data requirements because, as discussed in Chapter 2, male nuptiality data are not as readily available as female nuptiality data so obtaining estimates of the average age difference can be more difficult than obtaining data on the average age at marriage for females.

In conclusion, some negative association was identified between the average age difference and the indicators of women's status, as hypothesised. However, after controlling for the average age at marriage for females, little association was observed, which reflects the strong association between female age at marriage and women's status, at least at the aggregate level. Thus, associations between the age difference and indicators of women's status at the aggregate level are considered spurious due to the associations between the average age at marriage of females and in turn, the average age difference and women's status.

Chapter 4:
Results of examining cross-national and within-country variations in the age difference using individual-level data for 30 less developed countries

4.1 Overview

This chapter begins by using individual-level data to explore variations in the average age difference for 30 less developed countries, and variations in the distributions of the age difference within these countries. These analyses use data from the Demographic and Health Surveys to replicate the study by Casterline *et al* (1986) of World Fertility Survey data. However, several additions are made. Firstly, in addition to looking at the variations in the average age difference for all couples, this analysis also examines these variations just for monogamous couples since polygamous unions tend to have larger age differences on average (e.g. Dorjahn, 1959; Brown, 1981; Bongaarts *et al*, 1984; Garenne and van de Walle, 1989; Timaeus and Reynar, 1998). Given that the age difference is directly related to age at marriage (e.g. Dixon, 1971; Nawar, 1988; Bozon, 1991; Ní Bhrolcháin, 1992), variations in the average age at marriage for husbands and wives are examined cross-nationally and within countries. The association between the age difference and the age at marriage of husbands and then wives is then examined within-countries, initially for all couples but also just for monogamous couples. For each country, the observed distribution of age differences is then compared with the corresponding expected distribution. This is to test the hypothesis that variations in the average age difference and its distribution reflect the age-sex structure of those marrying such that mating occurs randomly with respect to age.

4.2 Cross-national variations in the average age difference

Table 4.1 presents descriptive statistics to summarise the distributions of the spousal age difference for couples in 30 less developed countries, which are categorised into five geographical regions according to the United Nations Statistics Division. Asterisks identify the eleven countries that were included in Casterline *et al*'s (1986) sample. As observed in Chapter 3, there is much variation in the average age difference between and within regions. Of this sample of countries, the median age difference ranges from 1.8 years (95% CI 1.5-2.0) in Kazakhstan to 9.7 years (95% CI 9.5-9.9) in Mali. Regional patterns are evident as

Table 4.1. Cross-national and within-country variation in the age difference: All couples

	Median age difference years, (95% C.I.)	Percentage distribution of the age difference <0	0-4	5-9	10-14	15+	Lower quartile	Upper quartile	Inter-quartile range	Number of couples (weighted, unweighted)
Central & South America										
Guatemala	2.1 (2.0-2.3)	13.4	54.8	22.4	6.2	3.2	0.1	5.1	5.0	2954, 3026
Peru *	2.4 (2.3-2.6)	14.8	48.9	25.1	7.6	3.6	0.0	5.6	5.6	7023, 7307
Brazil	2.9 (2.7-3.1)	16.7	42.5	27.1	8.0	5.6	0.2	6.4	6.3	2716, 2717
Colombia *	3.1 (2.9-3.4)	13.3	44.4	25.5	10.7	6.2	0.3	6.7	6.4	2500, 2523
Dominican Republic	3.9 (3.7-4.2)	7.9	42.7	28.9	12.4	8.1	1.3	7.7	6.4	1765, 1772
Haiti *	4.1 (3.6-4.6)	10.9	37.9	33.0	10.3	7.9	1.1	7.5	6.4	548, 540
Regional	*3.1*	*12.8*	*45.2*	*27.0*	*9.2*	*5.8*	*0.5*	*6.5*	*6.0*	
West Asia & North Africa										
Kazakhstan	1.8 (1.5-2.0)	10.9	66.4	19.1	3.1	0.5	0.0	3.8	3.8	933, 948
Uzbekistan	2.3 (2.1-2.5)	2.4	73.8	21.2	2.3	0.3	0.7	3.9	3.2	1383, 1361
Egypt *	5.8 (5.6-5.9)	2.6	30.7	42.6	18.5	5.6	2.9	8.9	6.0	5344, 5351
Regional	*3.3*	*5.3*	*57.0*	*27.6*	*8.0*	*2.1*	*1.2*	*5.5*	*4.3*	
South Asia										
Philippines *	2.1 (1.9-2.2)	17.5	50.4	22.8	6.3	3.0	-0.2	5.1	5.3	3650, 3694
Nepal *	2.9 (2.7-3.0)	6.9	55.9	28.5	5.9	2.7	1.0	5.3	4.3	3008, 2996
Pakistan *	4.4 (4.1-4.4)	5.4	39.9	35.0	12.8	6.9	1.9	7.7	5.8	2474, 2440
Bangladesh *	7.7 (7.5-7.9)	0.1	12.9	50.3	26.4	10.3	5.4	10.7	5.3	2551, 2580
Regional	*4.3*	*7.5*	*39.8*	*34.2*	*12.9*	*5.7*	*2.0*	*7.2*	*5.2*	
Central, East & South Africa										
Rwanda	3.8 (3.5-4.3)	11.8	39.9	31.9	10.7	5.6	0.8	7.1	6.3	1400, 1396
Uganda	4.3 (4.1-4.5)	3.7	43.1	35.9	10.9	6.4	2.0	7.4	5.4	1860, 1802
Malawi	4.5 (4.3-4.9)	1.1	42.4	40.0	10.4	6.1	2.5	7.7	5.3	997, 1065
Mozambique	4.8 (4.6-5.0)	3.4	39.5	31.4	14.8	10.9	2.3	9.1	6.9	2212, 2236
Central African Republic	4.9 (4.6-5.3)	3.9	38.8	31.4	15.0	10.8	2.2	9.2	7.0	1381, 1373
Zambia	4.9 (4.8-5.1)	0.8	37.5	44.5	12.5	4.8	3.0	7.6	4.6	2126, 2113
Zimbabwe	5.4 (5.1-5.7)	2.5	32.6	40.2	14.0	10.7	2.8	8.9	6.2	923, 877
Kenya *	5.4 (5.1-5.7)	2.0	31.9	42.4	14.2	9.6	3.1	8.8	5.7	1270, 1229
Tanzania	5.9 (5.4-6.2)	1.6	32.2	36.3	17.4	12.6	3.0	9.9	6.9	2211, 2148
Comoros	7.8 (6.9-8.6)	3.4	21.9	31.7	20.7	22.3	4.0	13.2	9.3	502, 502
Regional	*5.2*	*3.4*	*36.0*	*36.6*	*14.1*	*10.0*	*2.6*	*8.9*	*6.3*	
West Africa										
Ghana *	5.9 (5.5-6.4)	1.4	30.8	37.7	15.2	15.0	3.2	10.0	6.8	660, 660
Benin *	6.4 (6.1-6.7)	2.0	26.9	37.3	18.1	15.7	3.5	10.8	7.3	1615, 1626
Niger	8.6 (8.1-9.0)	1.5	14.5	37.3	26.2	20.6	4.1	12.9	8.8	1150, 1141
Burkina Faso	8.7 (8.4-9.1)	0.5	17.2	33.9	18.0	30.4	5.0	16.3	11.2	1757, 1642
Cameroon	9.1 (8.6-9.7)	0.6	18.6	29.5	25.1	26.3	4.8	14.3	9.6	683, 681
Cote d'Ivoire	9.2 (8.7-9.7)	2.3	15.4	31.3	22.5	28.4	5.3	15.1	9.8	1203, 1215
Mali	9.7 (9.5-9.9)	0.2	9.6	34.1	28.2	27.9	6.5	14.7	8.2	2661, 2629
Regional	*8.2*	*1.2*	*19.0*	*34.4*	*21.9*	*23.5*	*4.6*	*13.4*	*8.8*	

Notes for Table 4.1
* denotes that the country was included in Casterline *et al's* (1986) sample
Countries are presented in increasing order of age difference width, by region.
Regional statistics are not weighted & are simply the average of the national (weighted) estimates

Casterline *et al* (1986) noted. The average age difference is comparatively small for each of the six Central and South American samples, ranging from 2.1 years (95% CI 2.0-2.3) in Guatemala to 4.1 years (95% CI 3.6-4.6) in Haiti. As Casterline *et al* (1986) observed, the averages for the samples in this region are all less than five years. In contrast, the median age difference is greater than five years for all of the West African samples studied here and by Casterline *et al* (1986). Indeed, in this study's West African samples of Niger, Côte d'Ivoire, Cameroon, Burkina Faso, and Mali, half of couples have an age difference of at least eight years. Among the ten Central, East and South African countries considered by this study, age differences tend to be smaller on average than observed in the seven West African countries, with medians ranging from 3.8 years (95% CI 3.5-4.3) for Rwanda's sample to 7.8 years (95% CI 6.9-8.6) for the Comoros' sample. Indeed, excluding the Comoros, the next largest median for this regional sample is approximately two years smaller than the estimate for the Comoros, at 5.9 years (95% CI 5.4-6.2) for Tanzania's sample. Similar comparisons between West Africa and the rest of sub-Saharan Africa can not be made from Casterline *et al*'s (1986) study as they only studied five sub-Saharan African countries.

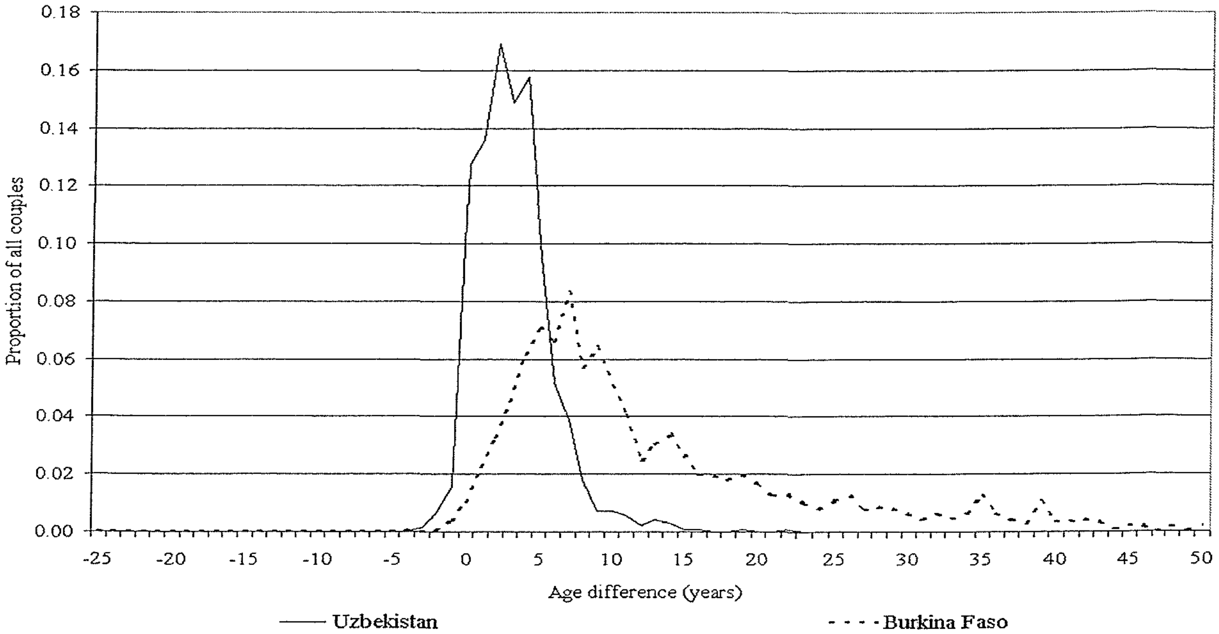
The other regions studied for this chapter display evidence of greater within-region variation. Among the four South Asian countries, the median age difference ranges from 2.1 years (95% CI 1.9-2.2) for the Philippines, which, as Casterline *et al* (1986) found, is smaller than any of the medians observed for the Central and South American countries, to 7.7 years (95% CI 7.5-7.9) for Bangladesh, which is larger than the medians observed for two of the West African countries. Indeed, the median age difference observed for the sample of couples in Bangladesh is more than three years greater than the sample mean for the four South Asian countries, 4.3 years. This observation can also be made from the estimates reported by Casterline *et al* (1986).

Within-region variation also exists for this study's West Asian and North African sample. For example, the median age difference observed for Egypt, 5.8 years (95% CI 5.6-5.9), is more than twice the size of the medians observed for the other two countries in this regional sample, Kazakhstan and Uzbekistan at 1.8 years (95% CI 1.5-2.0) and 2.3 years (95% CI 2.1-2.5) respectively. However, in the sample of six West Asian and North African countries that Casterline *et al* (1986) studied, the median for Egypt was estimated as 6.2 years, which was smaller than the unweighted regional mean of 7.1 years for their sample. While heterogeneity in the average age difference is observed for this region in both studies, it is important to note that their sample size for West Asia and North Africa is small. There is highly statistically significant regional variation in the average age difference (p=.000 according to an analysis of variance). Of the total variation in the average age difference for these 30 samples, 60.3% of this variation occurs between regions rather than within regions. The corresponding figures for Casterline *et al*'s (1986) analysis and the WISTAT3 analysis are similar and are calculated as 61.2% and 58.3%, respectively.

4.3 Variation in the age difference within countries

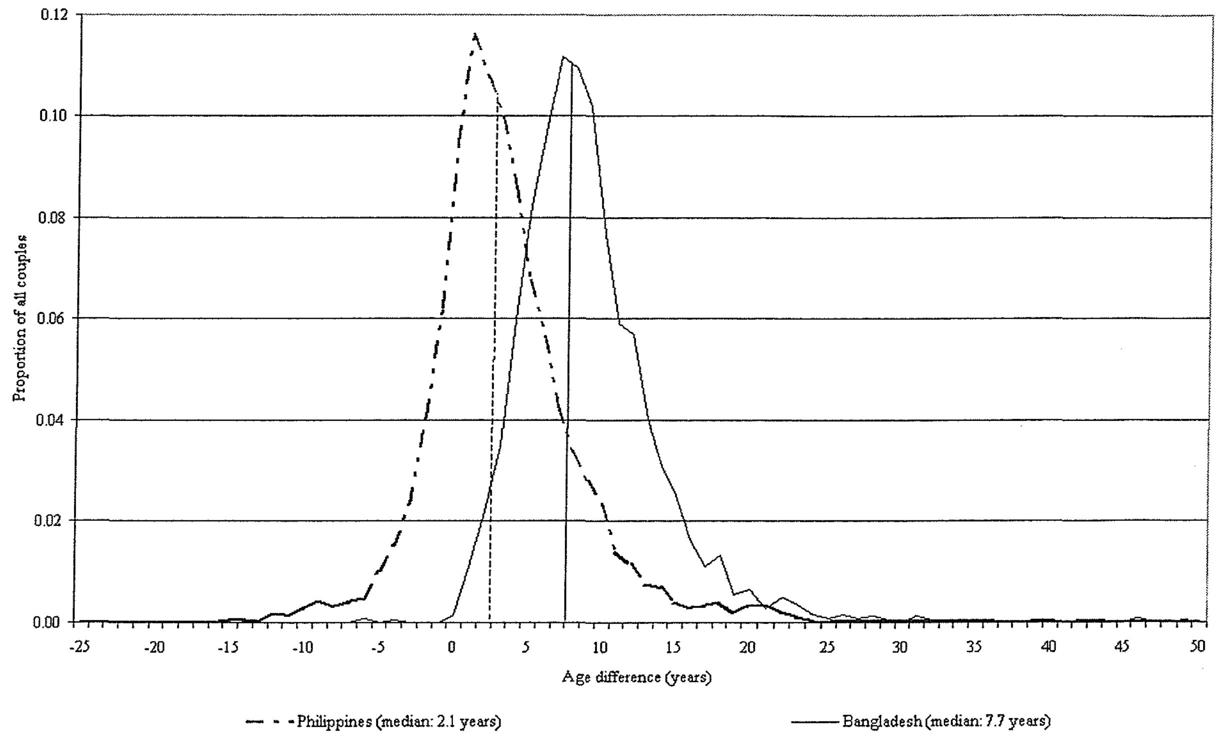
Of the total variation in the age difference that exists between the 61, 458 couples from the 30 countries studied, it is estimated from an analysis of variance that most, 83.5%, of the variation occurs within countries rather than between countries. Casterline *et al* (1986, p.368) reported a similar figure of 87.0%. This section now examines the variation in the distribution of age differences within countries as measured by the inter-quartile range (IQR). Table 4.1 shows that, of this sample of 30 countries, the smallest IQR is 3.2 years for Uzbekistan, i.e. half of the couples in this sample have an age difference between 0.7 years and 3.9 years. In contrast, the largest IQR is 11.2 years for Burkina Faso, i.e. half of the couples in this sample have an age difference between 5.0 years and 16.3 years. Of these 30 countries, Uzbekistan has the third smallest median age difference while Burkina Faso has the fourth largest median age difference. Figure 4.1 is a plot of the distribution of age differences for these two samples, and shows much more dispersion for Burkina Faso relative to Uzbekistan.

Figure 4.1 Distributions of age differences: Uzbekistan and Burkina Faso



In general, and as Casterline *et al* (1986) found, there is some positive association between the IQR and the median. The distribution of the age difference tends to be wider for countries with larger average age differences. However, this does not always apply. For example, the IQR is 5.3 years for both Bangladesh and the Philippines, however, there is 5.6 years difference between the medians for these two South Asian countries. Figure 4.2 shows the similarity in the dispersion of the distributions of the age difference for the Philippines and Bangladesh, despite the difference in the medians, which are highlighted.

Figure 4.2 Distributions of age differences: the Philippines and Bangladesh



Casterline *et al* (1986) also considered within-country variation in terms of the percentage distribution of the age difference, which they categorised into five groups: less than 0 years, 0-4 years, 5-9 years, 10-14 years, and at least 15 years. Table 4.1 uses these categories and shows that the largest proportion of any sample in one age difference category is 73.8% of couples in Uzbekistan having an age difference of between 0 and 4 years. Indeed, 95% of this sample have an age difference of between 0 and 9 years. This compares to Mali's sample, which has the smallest proportion of couples with an age difference of between 0 and 4 years, at 9.6%, and less than half of the sample, 43.7%, have an age difference between 0 and 9 years. Of the 28 countries studied by Casterline *et al* (1986) the largest percentage observed for any one such category is smaller at 52.5% but is again observed for age differences of between 0 and 4 years and for a country with a relatively small median age difference, Mexico, at 3.0 years. Another similarity with this study is that the country in Casterline *et al*'s (1986) study with the smallest percentage of its sample with an age difference of between 0 and 4 years, 7.8%, is a country with a relatively large median age difference, Bangladesh, at 9.1 years.

Looking now, as Casterline *et al* (1986) did, at the proportion of couples with more extreme age differences, such as where the wife is older than her husband, i.e. a negative age difference, and age differences of at least 15 years, variations in these proportions are evident both between and within these developing regions. Among the South Asian countries, is the sample with the smallest proportion of couples with a negative age difference of all 30 countries studied, Bangladesh (0.1%, which is similar to the figure of 0.2% reported for the WFS data by Casterline *et al*, 1986), and also the sample with the largest proportion of couples with a negative age difference of all 30 countries studied, the Philippines at 17.5%. While a similar proportion of negative age gaps (15.9%) was reported for the Philippines by Casterline *et al* (1986), Costa Rica was observed in their sample as having the largest proportion of negative age gaps at 17.7%. While Costa Rica is not among the 30 DHS surveys studied (since a DHS survey has not yet been conducted there), at least 10.9% of couples have a negative age difference in all but one of the seven Central and South American samples studied. The exception is the Dominican Republic where it is estimated that the wife is older than her husband in 7.9% of couples. Within-region variation in this measure is also observed for the four West Asian and North African countries. In Kazakhstan, 10.9% of couples have a negative age difference which compares to 2.4% and 2.6% of couples in Uzbekistan and Egypt, which is an interesting observation given that the medians for Uzbekistan and Kazakhstan are similar at 2.3 years and 1.8 years respectively. Among the sub-Saharan African countries studied, the proportion of couples with a negative age difference ranges from 0.2% for Mali to 11.8% for Rwanda, although excluding Rwanda the proportion is less than 4% of couples for all East, Central, South and West African samples. Rwanda seems quite unusual in this respect relative to the other sub-Saharan African countries studied by Casterline *et al* (1986), as the largest proportion of couples with a negative age difference that they report is 2.7% of couples in Benin.

Relative to the range in the proportion of couples with a negative age difference, there is greater variation between countries in terms of the proportion of couples with an age difference of at least 15 years. This ranges from 0.3% in the sample from Uzbekistan to 30.4% for the Burkina Faso sample, i.e. 30.1 percentage points relative to a range of 17.4 percentage points observed for the proportion of couples with a negative age difference. Among the West African samples, at least 15% of couples in this region's samples have an age difference of at least 15 years. Of the countries in other developing regional samples, the Comoros is the only country with a proportion of couples with an age difference of at least 15 years that is similar to the proportion observed for countries in West Africa at 22.3%. In the remaining 22 countries, the proportion is at most 12.6% of couples in Tanzania. Indeed, in many of these countries, the proportion is much smaller, i.e. less than 5%, and as Table 4.1 shows, there is some regional variation in this respect. For example, approximately 8% of couples in both the Dominican Republic and Haiti have an age difference of at least 15 years, which is a greater proportion than is observed for some sub-Saharan African countries such as Zambia, Rwanda, Malawi, and Uganda. These regional patterns were also observed by Casterline *et al* (1986). For example, while at most 8.4% of couples in their samples of Central and South American countries had an age difference of at least 15 years, in most of their sub-Saharan African samples between one fifth and one quarter of couples had such an age difference. An exception was Lesotho for which Casterline *et al* (1986) estimated that 9.8% of couples had an age difference of at least 15 years, which is smaller than they observed for several countries outside the sub-Saharan African region, e.g. Egypt (10.4%), Bangladesh (14.8%), and Pakistan (10.1%).

As reported by Casterline *et al* (1986), this section has identified much variation between countries in terms of the average age difference and in terms of the amount of variation in this difference within countries. The next section examines variations in the average age difference for monogamous unions in comparison to polygamous unions since it is well-established that age differences between a polygamous husband and his second or higher order wives tend to be larger, on average, than the age difference between a polygamous man and his first wife, or a monogamous man and his first wife (e.g. Dorjahn, 1959; Timaeus and Reynar, 1998).

4.4 Variations in the prevalence of polygamy and the association with the age difference

While Casterline *et al* (1986) mentioned polygamy as a possible explanation for the variations in the average age difference, evidence was not presented. Table 4.2 gives the median age difference and the 95% confidence interval for couples by marriage type (monogamous or polygamous) and the wife's rank relative to her co-wives for the DHS datasets studied that collected data on polygamy. Figure 4.3 then shows the variations in the distributions of the age difference by marriage type and wife's rank for the sub-Saharan African DHS datasets studied in this thesis. Figure 4.3 excludes Nepal and Pakistan, the two countries outside of sub-Saharan Africa in this study's sample in which questions on polygamy are asked, since less than 4% of women in these two samples are married to a polygamous husband (3.6% and 3.4% respectively, see Appendix 4A).

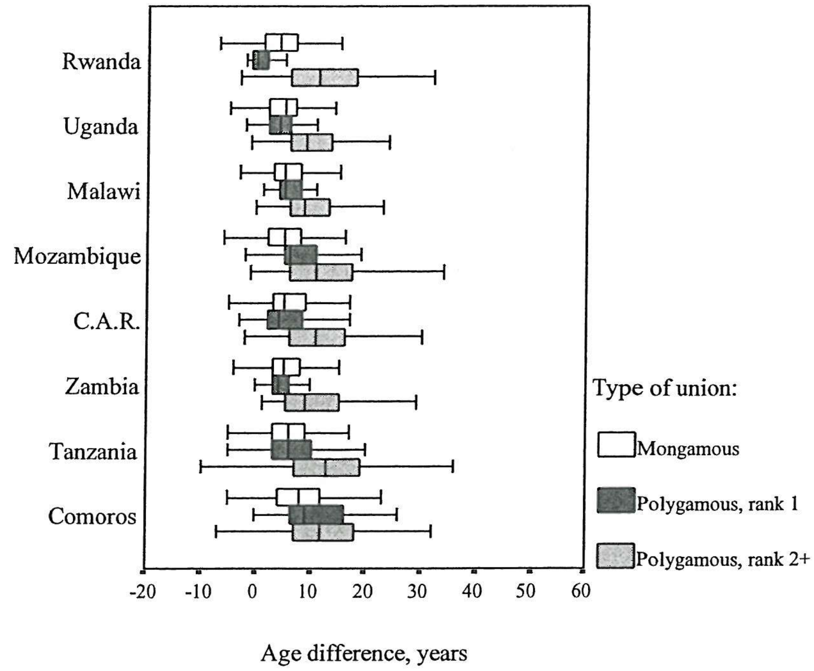
Table 4.2 Variations in the median age difference by union type and wife rank

Region and country	Median age difference (95% CI)		
Union type & wife rank:	Monogamous	Polygamous, rank 1	Polygamous, rank 2+
South Asia			
Nepal	2.8 (2.7-3.0)	0.7 (-0.7-3.1)	9.9 (7.1-12.6)
Pakistan	4.3 (4.1-4.4)	3.7 (0.8-12.4)	13.3 (8.9-17.3)
Central, East & South Africa			
Central African Republic	4.5 (4.2-4.9)	3.6 (2.9-4.7)	10.7 (9.3-11.7)
Comoros	7.1 (6.4-8.1)	7.8 (6.1-15.7)	11.6 (8.9-15.3)
Malawi	4.4 (4.1-4.7)	4.9 (3.8-7.2)	8.1 (6.8-10.2)
Mozambique	4.1 (3.9-4.4)	5.9 (5.3-7.4)	10.9 (9.2-12.1)
Rwanda	3.6 (3.3-4.0)	-0.3 (-1.8-2.8)	11.0 (9.1-15.2)
Tanzania	5.3 (5.0-5.6)	5.8 (4.3-7.5)	12.3 (10.5-13.4)
Uganda	4.0 (3.8-4.3)	3.1 (2.6-3.8)	8.3 (7.6-9.1)
Zambia	4.9 (4.7-5.1)	4.0 (3.1-4.2)	8.9 (6.6-10.4)
West Africa			
Benin	5.2 (4.9-5.6)	5.1 (4.5-5.9)	11.5 (10.3-12.6)
Burkina Faso	6.4 (6.2-6.7)	7.0 (6.3-8.4)	19.2 (18.3-20.4)
Cameroon	7.2 (6.4-8.0)	10.0 (6.5-11.7)	16.7 (15.5-19.4)
Côte d'Ivoire	7.5 (7.1-7.9)	8.8 (7.5-10.8)	18.2 (16.2-19.6)
Ghana	5.5 (5.0-6.0)	6.2 (5.0-8.1)	14.0 (11.3-17.2)
Mali	8.6 (8.3-8.9)	9.2 (8.1-9.9)	17.8 (16.9-18.8)
Niger	7.7 (7.4-8.1)	6.4 (5.0-9.1)	16.2 (14.5-18.0)

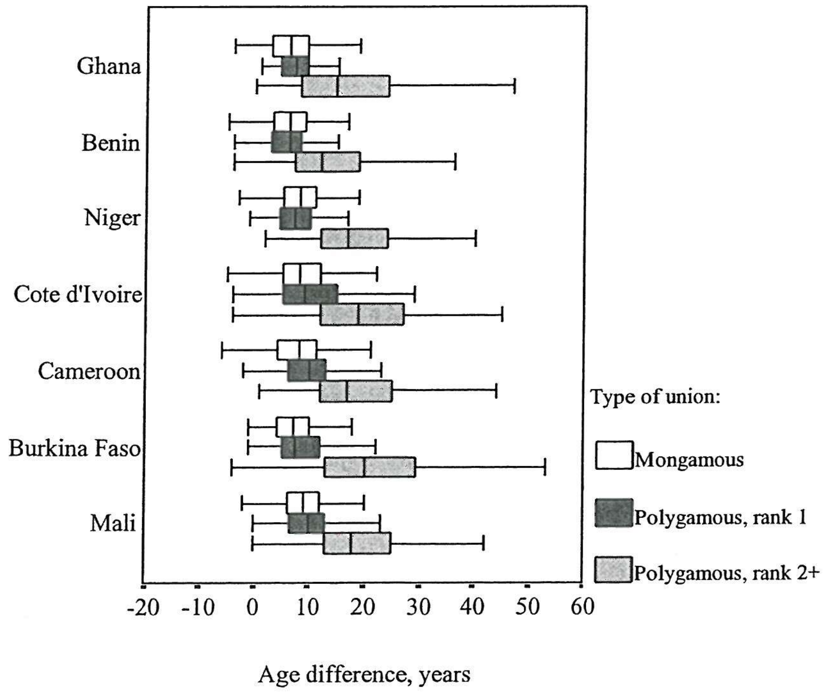
Note for Table 4.2
See Appendix 4A for percentage distribution of women by union type and wife's rank
Estimates are not presented for Kenya or Zimbabwe because these DHS surveys did not ask the wife's rank

Figure 4.3 Boxplots of the variation in the median age difference by union type and wife's rank for sub-Saharan African samples

Central, East and South Africa



West Africa



Notes for Figure 4.3
See Appendix 4A for percentage distribution of women by union type and wife's rank
Kenya or Zimbabwe are not presented because these DHS surveys did not ask about the wife's rank

To test for statistically significant variation in the average age difference by union type and wife rank, t-tests are used where variance homogeneity can be assumed, else Mann-Whitney tests are used. For all of the West African samples, there is no statistically significant variation in the mean age difference for women whose husbands were monogamous at the time of the interview in comparison to women who were their polygamous husband's first wife (p>.05). Among the Central, East and South African samples, the mean age difference in Mozambique is statistically significantly smaller for monogamous couples in comparison to women who were their polygamous husband's first wife (p=.000). In contrast, the mean age difference is statistically significantly *larger* for monogamous couples in comparison to women who were their polygamous husband's first wife for the samples of couples in Uganda and Zambia (p=.012 and p=.03, respectively). Comparing the age difference for polygamous couples by the wife's rank and for almost all samples, the average age difference is statistically significantly larger for each successive rank (p<.05).

Many of the samples that were identified in Table 4.1 with a relatively large median age difference and a large proportion of couples with an age difference of at least 15 years, are samples with a non-negligible proportion of women who were married to a polygamous husband and were a second or higher order wife (see Appendix 4A). For example, among the West African samples, between 10.9% in Ghana and 33.2% in Burkina Faso of women were married to a polygamous husband and were a second or higher order wife. The corresponding proportions were slightly smaller for the East, Central and South African countries, ranging between 4.6% of women in Rwanda and 15.3% of couples in Tanzania. Given the association between the age difference and the wife's rank, and the variation in the prevalence of polygamy in these 19 samples where such data were collected, the following sections briefly consider cross-national and within-country variations in the age difference solely among couples that were monogamous at interview for all 30 samples.

4.5 Cross-national variations in the average age difference among monogamous couples

Table 4.3 shows that, for surveys that collected data on polygamy, excluding women with polygamous husbands reduces the median age difference by as much as 2.3 years (for Burkina Faso). The magnitude of this change in the median age difference depends on the proportion of wives in the sample who were a second or higher order wife. Table 4.3 shows that, on

Table 4.3 Cross-national and within-country variation in the age difference: Monogamous couples only

	Median age difference years, (95% C.I.)	Percentage distribution of the age difference (years) <0	0-4	5-9	10-14	15+	Lower quartile	Upper quartile	Inter-quartile range	Number of monogamous couples (weighted, unweighted)	% of all couples monogamous
Central & South America											
Guatemala	2.1 (2.0-2.3)	13.4	54.8	22.4	6.2	3.2	0.1	5.1	5.0	2954, 3026	100.0
Peru	2.4 (2.3-2.6)	14.8	48.9	25.1	7.6	3.6	0.0	5.6	5.6	7023, 7307	100.0
Brazil	2.9 (2.7-3.1)	16.7	42.5	27.1	8.0	5.6	0.2	6.4	6.3	2716, 2717	100.0
Colombia	3.1 (2.9-3.4)	13.3	44.4	25.5	10.7	6.2	0.3	6.7	6.4	2500, 2523	100.0
Dominican Republic	3.9 (3.7-4.2)	7.9	42.7	28.9	12.4	8.1	1.3	7.7	6.4	1765, 1772	100.0
Haiti	4.1 (3.6-4.6)	10.9	37.9	33.0	10.3	7.9	1.1	7.5	6.4	548, 540	100.0
Regional	*3.1*	*12.8*	*45.2*	*27.0*	*9.2*	*5.8*	*0.5*	*6.5*	*6.0*		*100.0*
West Asia & North Africa											
Kazakhstan	1.8 (1.5-2.0)	10.9	66.4	19.1	3.1	0.5	0.0	3.8	3.8	933, 948	100.0
Uzbekistan	2.3 (2.1-2.5)	2.4	73.8	21.2	2.3	0.3	0.7	3.9	3.2	1383, 1361	100.0
Egypt	5.8 (5.6-5.9)	2.6	30.7	42.6	18.5	5.6	2.9	8.9	6.0	5344, 5351	100.0
Regional	*3.3*	*5.3*	*57.0*	*27.6*	*8.0*	*2.1*	*1.2*	*5.5*	*4.3*		*100.0*
South Asia											
Philippines	2.1 (1.9-2.2)	17.5	50.4	22.8	6.3	3.0	-0.2	5.1	5.3	3650, 3694	100.0
Nepal	2.8 (2.7-3.0)	6.9	56.9	28.7	5.5	2.0	1.0	5.2	4.2	2892, 2883	96.4
Pakistan	4.3 (4.1-4.4)	5.6	40.7	35.8	12.3	5.7	1.8	7.4	5.6	2383, 2333	96.6
Bangladesh	7.7 (7.5-7.9)	0.1	12.9	50.3	26.4	10.3	5.4	10.7	5.3	2551, 2581	100.0
Regional	*4.2*	*7.5*	*40.2*	*34.4*	*12.6*	*5.3*	*2.0*	*7.1*	*5.1*		*98.3*
Central, East & South Africa											
Rwanda	3.6 (3.3-4.0)	12.1	40.9	32.7	10.3	4.0	0.8	6.7	5.9	1322, 1319	93.4
Uganda	4.0 (3.8-4.3)	3.8	46.0	36.2	9.9	4.1	1.9	6.7	4.8	1484, 1433	78.0
Mozambique	4.1 (3.9-4.4)	4.0	44.7	30.4	13.5	7.3	2.0	8.0	6.0	1753, 1777	78.8
Malawi	4.4 (4.1-4.7)	1.3	44.6	39.7	9.5	4.9	2.3	7.4	5.1	898, 954	88.4
Central African Republic	4.5 (4.2-4.9)	4.3	41.2	33.6	12.8	8.0	2.0	8.0	6.0	1079, 1070	78.3
Zambia	4.9 (4.7-5.1)	0.8	37.9	45.3	12.2	3.9	3.0	7.4	4.4	1928, 1906	91.2
Kenya	5.0 (4.8-5.3)	2.1	34.3	45.0	13.9	4.7	2.9	7.9	5.0	1104, 1080	87.8
Zimbabwe	5.0 (4.8-5.4)	2.7	34.3	43.2	14.0	5.7	2.6	8.0	5.4	786, 737	85.9
Tanzania	5.3 (5.0-5.6)	1.7	35.6	39.3	16.0	7.4	2.7	8.7	6.0	1769, 1728	80.1
Comoros	7.1 (6.4-8.1)	3.6	24.3	32.2	21.2	18.8	3.6	11.9	8.3	416, 416	83.6
Regional	*4.8*	*3.6*	*38.4*	*37.8*	*13.3*	*6.9*	*2.4*	*8.1*	*6.1*		*84.6*
West Africa											
Benin	5.2 (4.9-5.6)	1.9	33.0	40.9	16.5	7.8	3.0	8.8	5.8	977, 968	60.6
Ghana	5.5 (5.0-6.0)	1.6	33.8	39.6	14.2	10.8	2.9	9.0	6.1	548, 548	81.0
Burkina Faso	6.4 (6.2-6.7)	0.3	25.2	46.6	18.7	9.2	3.9	9.6	5.7	1008, 1027	60.3
Cameroon	7.2 (6.4-8.0)	0.8	23.3	38.0	25.6	12.3	4.0	11.0	7.0	477, 503	73.3
Cote d'Ivoire	7.5 (7.1-7.9)	2.9	19.3	38.0	23.5	16.4	4.5	11.6	7.1	860, 867	72.9
Niger	7.7 (7.4-8.1)	1.8	16.7	42.2	26.2	13.2	4.7	10.9	6.2	916, 899	83.1
Mali	8.6 (8.3-8.9)	0.2	12.1	41.2	30.8	15.7	5.8	11.9	6.1	1886, 1872	71.1
Regional	*6.9*	*1.4*	*23.3*	*40.9*	*22.2*	*12.2*	*4.1*	*10.4*	*6.3*		*71.8*

Notes for Table 4.3
Countries are presented in increasing order of age difference width, by region.
Regional statistics are not weighted & are simply the average of the (weighted) national statistics

average, age differences still tend to be larger for monogamous couples in the West African samples relative to the other less developed regions. The smallest median observed for a West African sample of monogamous couples is 5.2 years (95% CI 4.9-5.6) for Benin, while the largest median is 8.6 years (95% CI 8.3-8.9) for monogamous couples in Mali. Among other regional samples there are only three other monogamous samples with a median age difference of a similar magnitude: Bangladesh at 7.7 years (95% CI 7.5-7.9), the Comoros at 7.1 years (95% CI 6.4-8.1), and Tanzania at 5.3 years (95% CI 5.0-5.6). In general therefore, average age differences tend to be larger among monogamous couples in West Africa than in other regions.

4.6 Within-country variations in the age difference among monogamous couples

Another consequence of focussing on monogamous couples is that the lower and upper quartiles are smaller in the distribution of age differences than when polygamous couples are included. Table 4.3 shows that the lower quartile is reduced by as much as 0.8 years for monogamous couples relative to all couples in Cameroon and Côte d'Ivoire. Reductions tend to be larger among the upper quartiles, especially among the West African samples where polygamy is more prevalent. Among the West African samples, the upper quartiles tend to be between two and three years smaller among monogamous couples relative to all couples. The largest difference is observed for Burkina Faso where the upper quartile is reduced from 16.3 years to 9.6 years by focussing on monogamous couples. Unsurprisingly, the proportion of couples with an age difference of at least 15 years is also greatly reduced relative to the sample of all couples. Again, the largest change is observed for Burkina Faso, where the proportion with an age difference of at least 15 years is reduced from 30.4% to 9.2%.

4.7 Variations in the age at marriage between and within countries

In addition to statistically significant variation in the age difference by union type, as discussed in Chapter 1, there is an association between the age difference and age at marriage. Thus, the next section examines cross-national and within-country variations in the average age at marriage for the 30 DHS samples for all couples and also just for monogamous couples, before going on to examine the association between the age difference and age at marriage.

4.7.1 Cross-national variations in the age at marriage of all husbands

Of the 30 samples of all couples, Table 4.4 shows that the median age at marriage for men whose wives married for the first time in the ten years prior to interview ranges from 19.9 years (95% CI 19.8-20.1) for Nepal –the only sample with a median of less than 20 years, to 28.2 years (95% CI 27.7-28.7) for Côte d'Ivoire. The median lies between 20.0 and 22.9 years in six samples: Guatemala (21.2 years, 95% CI 21.0-21.3), Kazakhstan (22.8 years, 95% CI 22.6-23.1) and Uzbekistan (22.2 years, 95% CI 22.0-22.4), and the three East, Central and South African countries of Uganda (22.0 years, 95% CI 21.8-22.3), Malawi (22.4 years, 95% CI 22.1-22.7), and Central African Republic (22.6 years, 95% CI 22.3-22.8). In over half of the samples, the median lies between ages 23.0 and 25.9. This includes four West African countries, six of the East, Central, and South African samples, all but one of the Central and South American samples (Guatemala), and all but Nepal in the South Asian sample. Thus, four of the six samples with a median age at marriage of at least 26 years are in West Africa, in addition to the Comoros (27.8 years, 95% CI 27.2-28.7) and Egypt (26.3 years, 95% CI 26.2-26.5). There is statistically significant regional variation in these patterns according to an analysis of variance (p=.002). However, as observed in Chapter 3 in the analysis of the aggregate data, more than half (55.2%) of the total variation in the average age at marriage of husbands occurs within rather than between these regional samples. Figure 4.4 plots the median age at marriage for husbands with the 95% confidence interval for each country, by region.

4.7.2 Cross-national variations in the age at marriage of monogamous husbands

Table 4.5 shows that of the 30 samples of monogamous couples, there is 7.2 years difference between the youngest median age at marriage of men, 19.9 years (95% CI 19.8-20.1) for Nepal, to 27.1 years (95% CI 26.4-27.8) for the sample of monogamous men in the Comoros. The oldest median is therefore 1.1 years younger than observed for the Côte d'Ivoire sample of all men, which had the oldest median age at marriage for all men. From comparing the unweighted regional mean age at marriage for monogamous men to that for all men in the Central, East and South African samples, it seems that monogamous men tend to marry 0.4 years earlier, on average, than the sample of all men. In comparison, monogamous men in the West African samples seem to marry on average 1.3 years earlier than all men. The difference between the West African and Central, East and South African samples reflects the greater prevalence of polygamy in West Africa, as observed above. Conversely, the low prevalence

Table 4.4 Cross-national and within-country variation in the age at marriage of husbands: All couples

	Median age at marriage, years (95% C.I.)	Percentage distribution of the husband's age at marriage (years) <20	20-22	23-25	26-29	30+	Lower quartile	Upper quartile	Inter-quartile range	Number of couples (weighted, unweighted)
Central & South America										
Guatemala	21.2 (21.0-21.3)	27.4	30.8	20.1	12.8	8.9	18.7	24.4	5.7	2954, 3026
Peru	23.4 (23.2-23.5)	12.5	25.9	23.4	18.8	19.4	20.5	27.4	7.0	7023, 7307
Brazil	23.9 (23.6-24.2)	11.1	24.7	22.9	20.3	20.9	20.9	27.9	7.1	2716, 2717
Dominican Republic	24.0 (23.7-24.4)	11.4	22.2	22.6	21.7	22.0	20.8	28.4	7.6	1765, 1772
Colombia	24.4 (24.2-24.7)	9.7	21.6	22.5	22.8	23.4	21.2	28.6	7.5	2500, 2523
Haiti	25.6 (25.0-26.1)	5.5	19.5	20.8	26.4	27.8	22.0	29.6	7.6	548, 540
Regional	*23.8*	*12.9*	*24.1*	*22.1*	*20.5*	*20.4*	*20.7*	*27.7*	*7.1*	
West Asia & North Africa										
Uzbekistan	22.2 (22.0-22.4)	7.5	39.1	35.8	13.1	4.4	20.8	24.1	3.3	1383, 1361
Kazakhstan	22.8 (22.6-23.1)	6.7	31.8	35.4	18.6	7.4	21.1	25.1	4.0	933, 948
Egypt	26.3 (26.2-26.5)	5.5	12.0	20.6	33.5	28.3	23.1	29.5	6.4	5344, 5351
Regional	*23.8*	*6.6*	*27.6*	*30.6*	*21.7*	*13.4*	*21.7*	*26.2*	*4.6*	
South Asia										
Nepal	19.9 (19.8-20.1)	37.6	31.6	16.6	8.0	6.1	17.9	22.8	4.9	3008, 2996
Bangladesh	23.4 (23.1-23.6)	11.6	26.2	24.4	23.6	14.2	20.7	26.8	6.1	2551, 2580
Philippines	23.5 (23.3-23.6)	11.8	26.4	25.2	20.7	15.9	20.5	26.8	6.3	3650, 3694
Pakistan	23.5 (23.3-23.8)	16.3	21.4	23.3	21.8	17.3	20.3	27.3	7.0	2474, 2440
Regional	*22.6*	*19.3*	*26.4*	*22.4*	*18.5*	*13.4*	*19.9*	*25.9*	*6.1*	
Central, East & South Africa										
Uganda	22.0 (21.8-22.3)	23.0	27.0	21.4	16.1	12.6	19.3	25.7	6.4	1860, 1802
Malawi	22.4 (22.1-22.7)	12.6	33.1	24.6	16.3	13.4	20.3	25.8	5.5	997, 1065
Central African Republic	22.6 (22.3-22.8)	21.4	23.6	19.9	16.9	18.2	19.4	27.1	7.6	1381, 1373
Mozambique	23.3 (23.1-23.6)	15.7	24.5	22.3	16.1	21.4	20.0	27.8	7.9	2212, 2236
Zambia	23.4 (23.2-23.7)	9.3	26.5	27.1	22.1	15.0	20.9	26.8	5.9	2126, 2113
Rwanda	24.7 (24.4-25.0)	6.8	19.3	26.2	27.4	20.3	21.8	27.8	6.0	1400, 1396
Zimbabwe	24.9 (24.5-25.4)	7.2	17.9	25.8	24.8	24.3	22.0	28.9	6.9	923, 877
Kenya	25.0 (24.8-25.3)	4.3	15.3	30.5	28.1	21.9	22.6	28.5	6.0	1270, 1229
Tanzania	25.1 (24.8-25.4)	6.1	18.6	24.7	24.0	26.6	22.0	29.5	7.5	2211, 2148
Comoros	27.8 (27.2-28.7)	6.2	10.4	16.1	23.3	44.0	23.5	33.9	10.4	502, 502
Regional	*24.1*	*11.3*	*21.6*	*23.9*	*21.5*	*21.8*	*21.2*	*28.2*	*7.0*	
West Africa										
Niger	24.0 (23.7-24.6)	15.2	17.9	22.9	19.0	25.0	20.9	29.0	8.2	1150, 1141
Benin	25.7 (25.3-26.1)	7.0	15.5	23.1	22.6	31.9	22.3	30.7	8.4	1615, 1626
Ghana	25.7 (25.1-26.3)	6.5	16.8	22.0	24.7	30.0	22.2	30.1	7.9	660, 660
Burkina Faso	26.4 (25.9-26.9)	6.7	14.5	19.7	20.4	38.6	22.5	33.9	11.3	1757, 1642
Mali	26.7 (26.4-27.0)	3.8	14.5	19.7	23.7	38.3	23.0	32.4	9.4	2661, 2629
Cameroon	27.0 (26.3-27.5)	7.0	12.3	16.4	27.8	36.4	22.7	31.7	9.0	683, 681
Cote d'Ivoire	28.2 (27.7-28.7)	5.5	10.7	15.8	22.5	45.5	23.8	33.8	10.0	1203, 1215
Regional	*26.2*	*7.4*	*14.6*	*19.9*	*23.0*	*35.1*	*22.5*	*31.7*	*9.2*	

Notes for Table 4.4
Countries are presented in increasing order of median age at marriage of husbands, by region.
Regional statistics are not weighted & are simply the average of the (weighted) national estimates

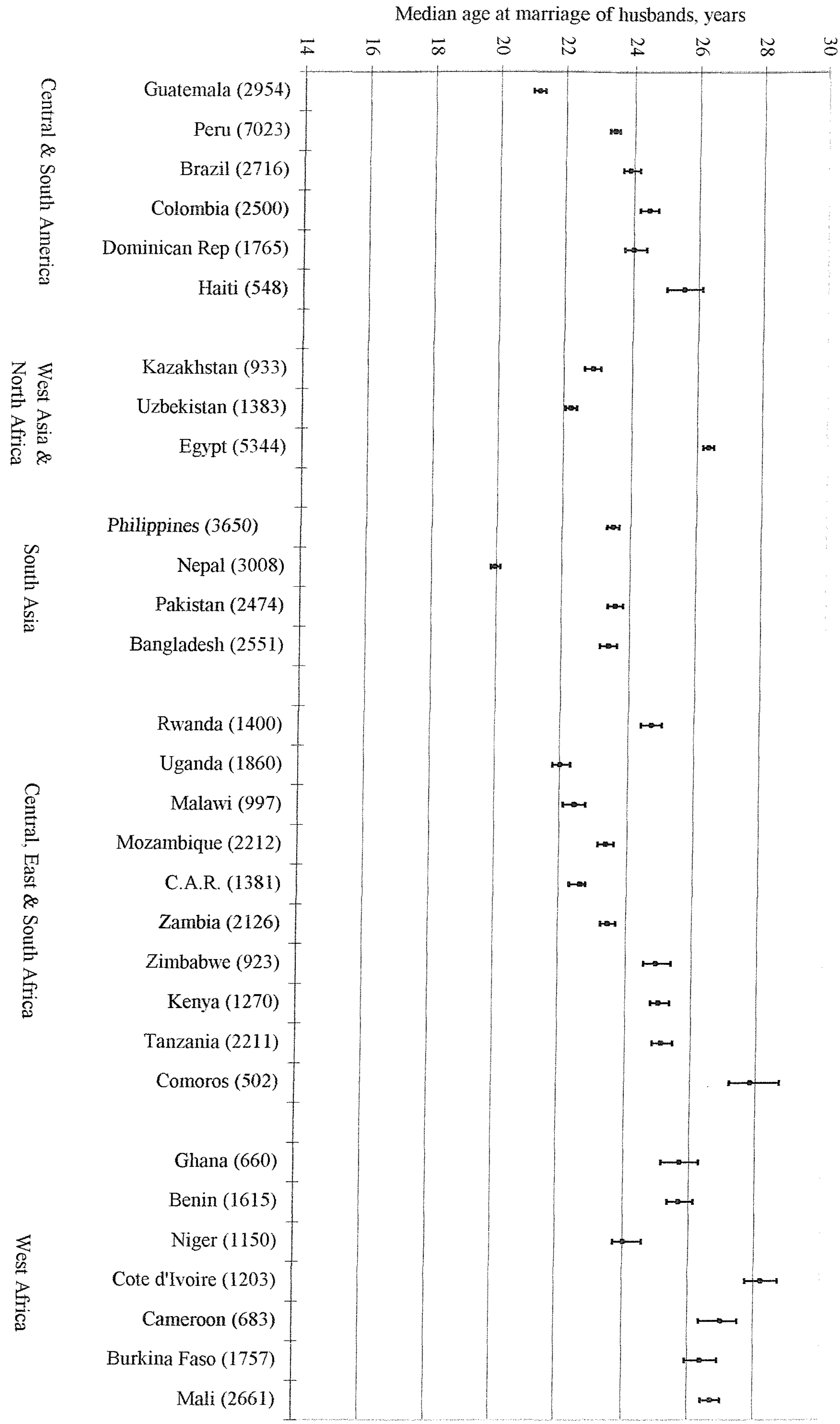


Figure 4.4: Median age at marriage of husbands (& 95% CI) by country and region (weighted sample size): All couples

Table 4.5 Cross-national and within-country variation in the age at marriage of husbands: Monogamous couples only

	Median age at marriage, years (95% C.I.)	Percentage distribution of the husband's age at marriage (years) <20	20-22	23-25	26-29	30+	Lower quartile	Upper quartile	Inter-quartile range	Number of monogamous couples (weighted, unweighted)
Central & South America										
Guatemala	21.2 (21.0-21.3)	27.4	30.8	20.1	12.8	8.9	18.7	24.4	5.7	2954, 3026
Peru	23.4 (23.2-23.5)	12.5	25.9	23.4	18.8	19.4	20.5	27.4	7.0	7023, 7307
Brazil	23.9 (23.6-24.2)	11.1	24.7	22.9	20.3	20.9	20.9	27.9	7.1	2716, 2717
Dominican Republic	24.0 (23.7-24.4)	11.4	22.2	22.6	21.7	22.0	20.8	28.4	7.6	1765, 1772
Colombia	24.5 (24.2-24.7)	9.7	21.6	22.5	22.8	23.4	21.2	28.6	7.5	2500, 2523
Haiti	25.6 (25.0-26.1)	5.5	19.5	20.8	26.4	27.8	22.0	29.6	7.6	548, 540
Regional	*23.8*	*12.9*	*24.1*	*22.1*	*20.5*	*20.4*	*20.7*	*27.7*	*7.1*	
West Asia & North Africa										
Uzbekistan	22.2 (22.0-22.4)	7.5	39.1	35.8	13.1	4.4	20.8	24.1	3.3	1383, 1361
Kazakhstan	22.8 (22.6-23.1)	6.7	31.8	35.4	18.6	7.4	21.1	25.1	4.0	933, 948
Egypt	26.3 (26.2-26.5)	5.5	12.0	20.6	33.5	28.3	23.1	29.5	6.4	5344, 5351
Regional	*23.8*	*6.6*	*27.6*	*30.6*	*21.7*	*13.4*	*21.7*	*26.2*	*4.6*	
South Asia										
Nepal	19.9 (19.7-20.0)	38.4	32.0	16.7	7.7	5.2	17.8	22.6	4.8	2892, 2883
Bangladesh	23.4 (23.1-23.6)	11.6	26.2	24.4	23.6	14.2	20.7	26.8	6.1	2551, 2580
Pakistan	23.4 (23.2-23.6)	16.6	21.9	23.8	21.7	16.7	20.2	27.1	6.9	2383, 2333
Philippines	23.5 (23.3-23.6)	11.8	26.4	25.2	20.7	15.9	20.5	26.8	6.3	3650, 3694
Regional	*22.5*	*19.6*	*26.6*	*22.5*	*18.4*	*13.0*	*19.8*	*25.8*	*6.0*	
Central, East & South Africa										
Uganda	21.7 (21.5-22.0)	24.5	28.5	22.6	15.6	8.8	19.1	24.9	5.8	1484, 1433
Central African Republic	22.1 (21.8-22.5)	23.6	25.2	21.0	16.3	13.9	19.2	26.1	6.9	1079, 1070
Malawi	22.2 (21.9-22.5)	13.3	34.5	25.4	15.5	11.3	20.2	25.3	5.1	898, 954
Mozambique	22.6 (22.2-23.0)	18.1	28.3	21.4	15.3	17.0	19.6	26.5	6.9	1753, 1777
Zambia	23.4 (23.2-23.6)	9.2	26.5	28.0	22.4	13.8	20.9	26.6	5.7	1928, 1906
Tanzania	24.4 (24.1-24.7)	6.7	20.5	27.2	24.7	20.9	21.8	27.7	6.0	1769, 1728
Zimbabwe	24.5 (24.2-24.8)	7.6	18.9	28.6	25.4	19.5	21.8	27.9	6.1	786, 737
Rwanda	24.5 (23.8-24.7)	7.0	18.8	27.1	28.0	18.2	21.8	27.5	5.8	1322, 1319
Kenya	24.7 (24.4-24.9)	4.6	16.6	33.0	29.0	16.8	22.4	27.7	5.4	1104, 1080
Comoros	27.1 (26.4-27.8)	7.5	11.3	17.8	25.0	38.5	23.0	33.0	10.0	416, 416
Regional	*23.7*	*12.2*	*22.9*	*25.2*	*21.7*	*17.9*	*21.0*	*27.3*	*6.4*	
West Africa										
Niger	23.2 (22.9-23.6)	17.0	20.9	26.4	19.5	16.2	20.4	26.7	6.3	916, 899
Burkina Faso	24.1 (23.7-24.5)	10.1	20.1	27.4	24.1	18.3	21.4	27.6	6.2	1008, 1027
Benin	24.7 (24.3-25.1)	8.5	17.9	26.3	23.4	24.0	21.8	28.8	7.0	977, 968
Ghana	25.1 (24.5-25.8)	7.7	17.9	23.5	25.9	25.0	21.9	29.0	7.1	548, 548
Cameroon	25.4 (24.9-25.9)	10.8	15.7	19.7	29.4	24.5	21.7	28.9	7.2	477, 503
Mali	25.4 (25.2-25.7)	4.9	17.5	23.9	26.8	26.9	22.3	27.6	5.3	1886, 1872
Cote d'Ivoire	26.6 (26.1-27.0)	6.9	13.8	19.2	25.1	35.1	23.0	31.0	8.1	860, 867
Regional	*24.9*	*9.4*	*17.7*	*23.8*	*24.9*	*24.3*	*21.8*	*28.5*	*6.7*	

Notes for Table 4.5
Countries are presented in increasing order of median age at marriage of husbands, by region.
Regional statistics are not weighted & are simply the average of the (weighted) national estimates

of polygamy in the two South Asian countries in this sample results in little difference between the median age at marriage for monogamous men and that of all men in the samples of Nepal and Pakistan. Overall, there is no statistically significant regional variation in the average age at marriage for monogamous men according to an analysis of variance (p=.199). More than three-quarters (79.4%) of the total variation occurs within rather than between these regional samples of monogamous men.

4.7.3 Cross-national variations in the age at marriage of all wives

Table 4.6 shows that of the 30 samples studied, the median age at first marriage of women who married in the ten years prior to interview ranges from 14.5 years (95% CI 14.4-14.8) for Bangladesh to 21.2 years (95% CI 21.1-21.4) for the Philippines. In addition to Bangladesh, the median is also less than 16 years in two other countries: Niger and Mali. The median lies between ages 16.0 and 18.9 in over half of the samples, which includes the remaining five West African countries, all but one of the East, Central, and South African samples (Rwanda), two of the Central and South American samples (Guatemala and the Dominican Republic), and Pakistan in the South Asian sample. Thus, the samples with a median age at marriage of at least 19 years are mainly in Central and South America, and West Asia and North Africa. Overall, there is statistically significant regional variation in the average age at marriage of wives according to an analysis of variance (p=.004). However, a slightly larger proportion (55.5%) of the total variation occurs within rather than between these regional samples. Figure 4.5 plots the median age at marriage for wives with the 95% confidence interval for each country, by region.

4.7.4 Cross-national variations in the age at marriage of wives of monogamous men

Comparing Tables 4.6 and 4.7, there is virtually no difference between the average age at marriage of wives married to monogamous husbands and the average age at marriage of all wives. Among the West African samples, where polygamy is most prevalent, there is 0.1 years difference, on average, between medians for all couples in comparison to just monogamous couples. Looking at the difference for individual countries, there is no consistency in the direction of the difference, i.e. for some samples, the median is slightly larger for monogamous couples only, while for some samples, the median is slightly smaller for monogamous couples only. Thus, the same observations can be made for the samples of monogamous couples that were described above for the samples of all couples. Overall, there

Table 4.6 Cross-national and within-country variation in the age at first marriage of wives: All couples

	Median age at marriage, years (95% C.I.)	Percentage distribution of the wife's age at first marriage (years) <16	16-18	19-21	22-24	25+	Lower quartile	Upper quartile	Inter-quartile range	Number of couples (weighted, unweighted)
Central & South America										
Guatemala	17.9 (17.8-18.1)	14.2	36.5	27.3	12.1	9.9	15.9	20.6	4.7	2954, 3026
Dominican Republic	18.9 (18.7-19.2)	13.8	27.5	27.1	14.5	17.1	16.3	22.2	5.9	1765, 1772
Haiti	20.3 (17.6-24.1)	5.6	23.3	25.9	19.0	26.0	17.6	24.1	6.5	548, 540
Peru	20.0 (19.9-20.1)	6.8	23.4	28.3	18.2	23.3	17.5	23.7	6.2	7023, 7307
Colombia	20.2 (20.0-20.5)	6.1	22.5	28.1	19.6	23.7	17.6	23.7	6.1	2500, 2523
Brazil	20.3 (20.1-20.5)	6.8	22.5	27.6	19.6	23.5	17.5	23.7	6.2	2716, 2717
Regional	*19.6*	*8.9*	*26.0*	*27.4*	*17.2*	*20.6*	*17.1*	*23.0*	*5.9*	
West Asia & North Africa										
Uzbekistan	19.3 (19.1-19.4)	0.6	26.0	50.1	16.9	6.4	17.9	20.9	3.0	1383, 1361
Egypt	19.4 (19.3-19.6)	8.6	27.9	26.6	21.0	15.8	16.8	22.5	5.7	5344, 5351
Kazakhstan	20.4 (20.2-20.6)	0.8	18.5	40.8	27.9	12.0	18.4	22.4	3.9	933, 948
Regional	*19.5*	*3.3*	*24.1*	*39.2*	*21.9*	*11.4*	*17.7*	*21.9*	*4.2*	
South Asia										
Bangladesh	14.5 (14.4-14.8)	55.5	30.9	10.0	2.7	0.9	13.3	16.7	3.4	2551, 2580
Nepal	16.3 (16.2-16.4)	28.4	44.1	18.7	5.9	2.8	14.8	18.3	3.5	3008, 2996
Pakistan	18.0 (17.8-18.1)	16.9	33.7	26.3	13.6	9.4	15.8	20.7	5.0	2474, 2440
Philippines	21.2 (21.1-21.4)	4.5	22.2	30.0	20.6	22.6	18.8	24.5	5.7	3650, 3694
Regional	*17.5*	*26.3*	*32.7*	*21.3*	*10.7*	*8.9*	*15.7*	*20.1*	*4.4*	
Central, East & South Africa										
Central African Republic	16.59 (16.4-16.8)	26.5	42.6	19.6	6.1	5.2	14.9	18.7	3.8	1381, 1373
Uganda	16.9 (16.7-17.0)	19.6	46.8	23.2	7.0	3.4	15.3	18.9	3.5	1860, 1802
Mozambique	17.0 (16.9-17.2)	18.6	43.0	22.8	9.5	6.0	15.5	19.4	3.9	2212, 2236
Malawi	17.2 (17.0-17.4)	16.4	47.3	26.3	6.8	3.2	15.7	18.8	3.1	997, 1065
Zambia	17.6 (17.5-17.7)	10.3	47.0	27.2	10.2	5.3	16.1	19.7	3.6	2126, 2113
Tanzania	18.2 (18.1-18.4)	7.9	39.2	32.2	12.9	7.9	16.6	20.5	3.9	2211, 2148
Zimbabwe	18.6 (18.4-18.9)	9.0	33.6	33.9	15.9	7.6	16.7	20.9	4.2	923, 877
Kenya	18.7 (18.5-18.9)	8.4	32.9	33.5	15.8	9.4	16.7	21.0	4.3	1270, 1229
Comoros	18.9 (18.5-19.4)	11.6	28.3	27.1	15.9	17.1	16.5	22.4	5.9	502, 502
Rwanda	20.0 (19.8-20.3)	2.9	20.5	36.8	26.2	13.6	18.1	22.5	4.4	1400, 1396
Regional	*18.0*	*13.1*	*38.1*	*28.3*	*12.6*	*7.9*	*16.2*	*20.3*	*4.1*	
West Africa										
Niger	14.9 (14.7-15.0)	53.4	35.5	7.6	2.4	1.1	13.6	16.2	2.7	1150, 1141
Mali	15.7 (15.7-15.8)	34.8	44.3	13.2	4.4	3.3	14.6	17.6	3.0	2661, 2629
Cameroon	16.4 (16.1-16.7)	26.4	38.5	21.0	8.1	6.0	14.6	18.7	4.1	683, 681
Burkina Faso	16.7 (16.6-16.8)	16.4	57.2	22.2	2.7	1.4	15.6	18.1	2.5	1757, 1642
Cote d'Ivoire	17.6 (17.4-17.8)	16.5	38.7	24.7	11.1	9.1	15.8	20.1	4.4	1203, 1215
Benin	18.1 (17.9-18.2)	9.5	39.7	31.4	11.4	8.0	16.4	20.3	3.8	1615, 1626
Ghana	18.6 (18.3-18.9)	7.4	34.2	33.9	15.6	8.8	16.6	20.9	4.3	660, 660
Regional	*16.9*	*23.5*	*41.2*	*22.0*	*8.0*	*5.4*	*15.3*	*18.8*	*3.5*	

Notes for Table 4.6

Countries are presented in increasing order of median age at marriage of wives, by region.

Regional statistics are not weighted & are simply the average of the (weighted) national estimates

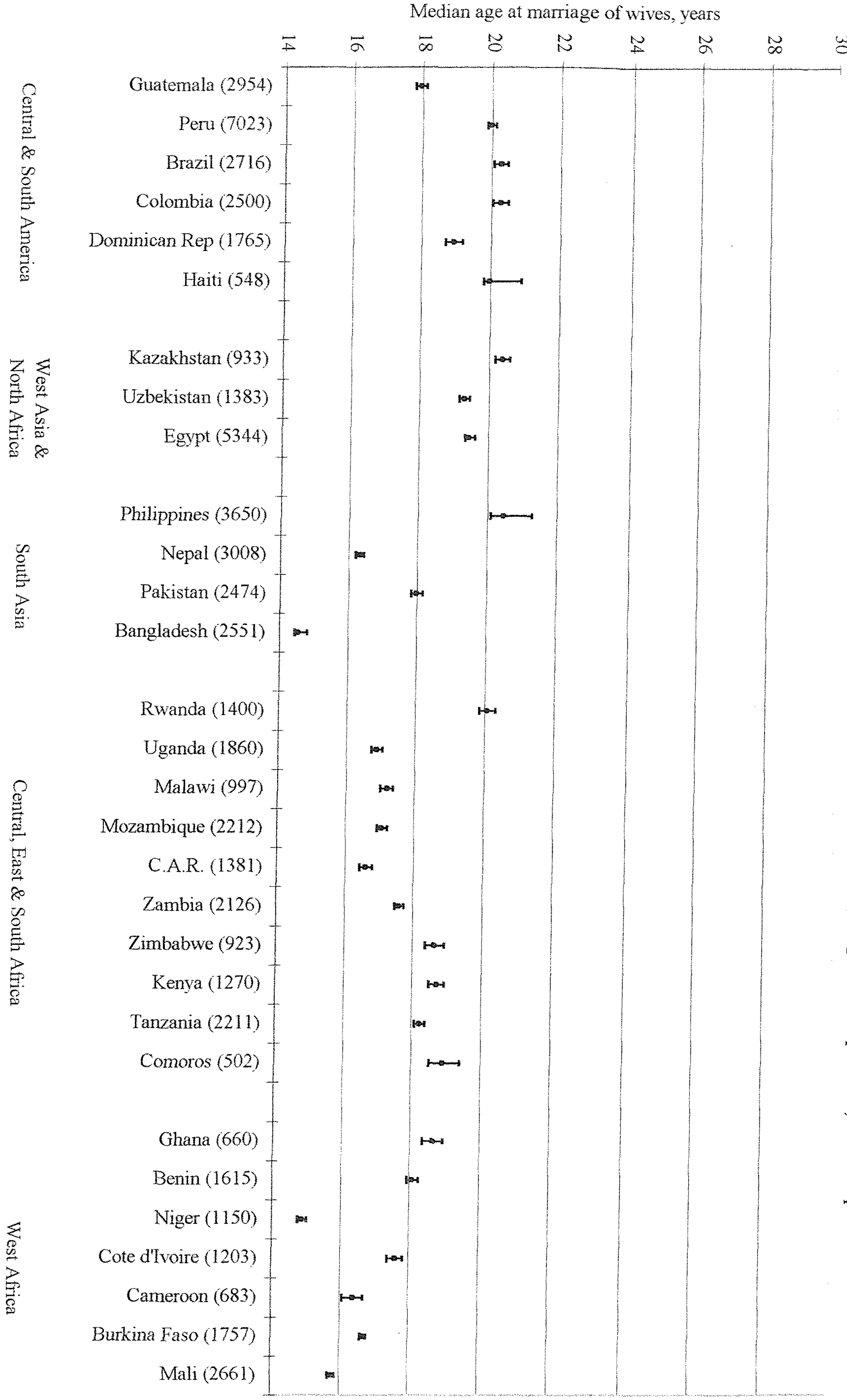


Figure 4.5: Median age at marriage of wives (& 95% CI) by country and region (weighted sample size): All couples

Table 4.7 Cross-national and within-country variation in the age at first marriage of wives: Monogamous couples only

	Median age at marriage, years (95% C.I.)	Percentage distribution of the wife's age at first marriage (years) <16	16-18	19-21	22-24	25+	Lower quartile	Upper quartile	Inter-quartile range	Number of monogamous couples (weighted, unweighted)
Central & South America										
Guatemala	17.9 (17.8-18.1)	14.2	36.5	27.3	12.1	9.9	15.9	20.6	4.7	2954, 3026
Dominican Republic	18.9 (18.7-19.2)	13.8	27.5	27.1	14.5	17.1	16.3	22.2	5.9	1765, 1772
Haiti	20.3 (17.6-24.1)	5.6	23.3	25.9	19.0	26.0	17.6	24.1	6.5	548, 540
Peru	20.0 (19.9-20.1)	6.8	23.4	28.3	18.2	23.3	17.5	23.7	6.2	7023, 7307
Colombia	20.2 (20.0-20.5)	6.1	22.5	28.1	19.6	23.7	17.6	23.7	6.1	2500, 2523
Brazil	20.3 (20.1-20.5)	6.8	22.5	27.6	19.6	23.5	17.5	23.7	6.2	2716, 2717
Regional	*19.6*	*8.9*	*26.0*	*27.4*	*17.2*	*20.6*	*17.1*	*23.0*	*5.9*	
West Asia & North Africa										
Uzbekistan	19.3 (19.1-19.4)	0.6	26.0	50.1	16.9	6.4	17.9	20.9	3.0	1383, 1361
Egypt	19.4 (19.3-19.6)	8.6	27.9	26.6	21.0	15.8	16.8	22.5	5.7	5344, 5351
Kazakhstan	20.4 (20.2-20.6)	0.8	18.5	40.8	27.9	12.0	18.4	22.4	3.9	933, 948
Regional	*19.5*	*3.3*	*24.1*	*39.2*	*21.9*	*11.4*	*17.7*	*21.9*	*4.2*	
South Asia										
Bangladesh	14.5 (14.4-14.8)	55.5	30.9	10.0	2.7	0.9	13.3	16.7	3.4	2551, 2580
Nepal	16.3 (16.2-16.4)	28.4	44.4	18.8	5.8	2.6	14.8	18.3	3.5	2892, 2883
Pakistan	18.0 (17.8-18.2)	16.6	34.0	26.5	13.7	9.2	15.8	20.7	4.9	2383, 2333
Philippines	21.2 (21.1-21.4)	4.5	22.2	30.0	20.6	22.6	18.8	24.5	5.7	3650, 3694
Regional	*17.5*	*26.3*	*32.9*	*21.3*	*10.7*	*8.8*	*15.7*	*20.1*	*4.4*	
Central, East & South Africa										
Central African Republic	16.5 (16.3-16.8)	25.9	44.9	18.9	5.8	4.6	14.9	18.5	3.6	1079, 1070
Uganda	16.8 (16.6-16.9)	19.4	48.3	23.0	6.6	2.7	15.3	18.7	3.4	1484, 1433
Mozambique	17.1 (16.9-17.3)	17.4	44.3	22.2	9.3	6.7	15.6	19.4	3.8	1753, 1777
Malawi	17.1 (16.9-17.3)	16.6	47.3	26.6	6.5	3.0	15.7	18.8	3.1	898, 954
Zambia	17.6 (17.4-17.7)	10.3	47.1	27.4	9.9	5.2	16.1	19.6	3.6	1928, 1906
Tanzania	18.2 (18.0-18.4)	7.3	39.4	32.8	13.1	7.3	16.6	20.5	3.9	1769, 1728
Kenya	18.7 (18.4-18.9)	8.3	33.4	33.6	15.5	9.1	16.7	20.9	4.3	1104, 1080
Zimbabwe	18.7 (18.4-19.0)	9.1	32.8	34.1	16.9	7.1	16.7	20.9	4.2	786, 737
Comoros	18.7 (18.3-19.1)	12.5	29.3	27.9	16.1	14.2	16.3	22.0	5.7	416, 416
Rwanda	20.0 (19.8-20.3)	3.0	20.3	37.2	26.1	13.5	18.1	22.5	4.3	1322, 1319
Regional	*17.9*	*13.0*	*38.7*	*28.4*	*12.6*	*7.3*	*16.2*	*20.2*	*4.0*	
West Africa										
Niger	14.8 (14.6-14.9)	55.5	34.2	7.2	2.2	0.8	13.6	16.0	2.4	916, 899
Mali	15.8 (15.7-15.9)	34.2	44.8	13.5	4.2	3.3	14.6	17.6	3.0	1886, 1872
Cameroon	16.7 (16.3-17.1)	29.0	37.4	21.0	7.5	5.1	14.7	19.0	4.3	477, 503
Burkina Faso	16.7 (16.6-16.8)	16.2	54.5	23.7	4.0	1.7	15.5	18.3	2.8	1008, 1027
Cote d'Ivoire	17.7 (17.4-18.0)	17.0	36.9	24.1	12.1	10.0	15.8	20.4	4.6	860, 867
Benin	18.2 (17.9-18.4)	8.7	39.2	31.4	12.1	8.6	16.5	20.5	3.9	977, 968
Ghana	18.6 (18.3-18.9)	7.7	33.9	33.8	16.8	7.8	16.6	21.0	4.3	548, 548
Regional	*16.9*	*24.0*	*40.1*	*22.1*	*8.4*	*5.3*	*15.3*	*19.0*	*3.6*	

Notes for Table 4.7

Countries are presented in increasing order of median age at marriage of wives, by region.

Regional statistics are not weighted & are simply the average of the (weighted) national estimates

is statistically significant regional variation in the average age at marriage of women with monogamous husbands according to an analysis of variance (*p*=.006). Similar to the sample of all women, a slightly larger proportion (57.3%) of the total variation occurs within rather than between these regional samples.

4.7.5 Within-country variations in the age at marriage of all husbands

The inter-quartile range (IQR) is used as a measure of within-country variation in the age at marriage of husbands. Table 4.4 shows that this ranges from 3.3 years for Uzbekistan to 11.3 years for Burkina Faso. Thus, half of the men in the sample of Uzbekistan married between ages 20.8 years and 24.1 years, while half of the men in the sample of Burkina Faso married between 22.5 years and 33.9 years. IQRs of the magnitude observed for Burkina Faso are not unusual among the West African samples, with the unweighted regional mean IQR at 9.2 years. The IQR for the Comoros sample is of a similar magnitude at 10.4 years, although generally, IQRs tend to be a couple of years smaller for Central, East and South African samples, with an unweighted regional mean IQR of 7.0 years, which is similar to that calculated for the Central and South American samples. Overall, there is statistically significant variation between countries in the husband's age at marriage (*p*<.05 according to a Kruskal-Wallis test as variance homogeneity can not be assumed). However, 91.0% of the total variation occurs within rather than between countries.

4.7.6 Within-country variations in the age at marriage of monogamous husbands

Table 4.5 shows that there is less within-sample variation in the average age at marriage of monogamous husbands as measured by the IQR than for the samples of all husbands. Among the West African samples, this tends to be at least a couple of years narrower on average. The greatest difference in the IQR between the sample of all couples and the sample of monogamous couples only is observed for Burkina Faso, where the IQR is reduced by 5.1 years by focussing on monogamous men. There is also less within-sample variation among the Central, East and South African samples albeit to a lesser degree. On average, the IQR is reduced by 0.6 years by focussing on monogamous men. As might be expected, there is greater difference between the upper quartiles when comparing the samples of all men with the samples of just monogamous men, relative to the lower quartiles. Indeed, by focussing just on monogamous men, the upper quartiles for the sub-Saharan African samples are more similar to those observed for the samples that did not ask about polygamy, especially in

Central and South America. This is also evident when looking at the percentage of men who married aged at least 30. In four of the Central and South American samples and in five of the West African samples, at least one in five men married aged at least 30. According to an analysis of variance, there is highly statistically significant variation between countries in the age at marriage of monogamous husbands (p=.000). However, 93.2% of the total variation occurs within rather than between countries. Figure 4.6 presents boxplots to show within-country variations in the husband's age at marriage by union type for countries in sub-Saharan Africa. The greater variation in the age at marriage of polygamous men married to a wife of at least rank two relative to monogamous men and polygamous men married to a wife of rank one is clearly visible for all countries.

Figure 4.6 Boxplots to show within-country variations in the husband's age at marriage by union type for countries in sub-Saharan Africa

Central, East and South Africa

West Africa

Notes for Figure 4.6
See Appendix 4A for percentage distribution of women by union type and wife's rank
Kenya or Zimbabwe are not presented because these DHS surveys did not ask about the wife's rank

4.7.7 Within-country variations in the age at marriage of all wives

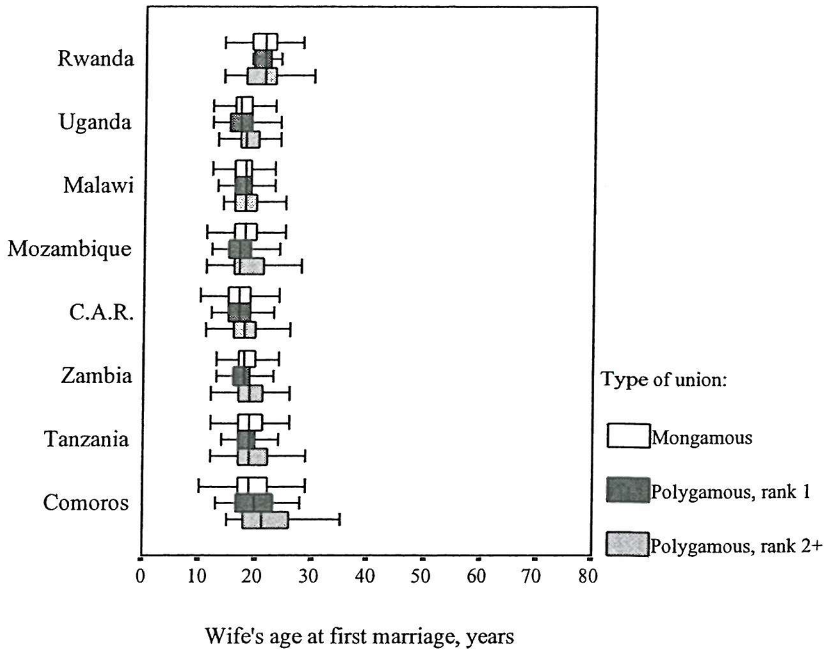
The IQR is again used to measure within-country variation in the average age at marriage of wives. Table 4.6 shows that for all countries, it is smaller than the corresponding IQR for the age at marriage of husbands. The IQR ranges from 2.5 years for the sample of women in Burkina Faso to 6.5 years in Haiti. Thus, half of women in the sample of Burkina Faso married between ages 15.6 and 18.1 years, while half of women in the sample of Haiti married between ages 17.6 years and 24.1 years. In general, there is a positive association between the IQR and the median age at marriage for women, such that in samples where the median is relatively early, the IQR is relatively small. For example, the regional average for the Central and South American samples is 19.6 years and the regional IQR is 5.9 years, which compares to 16.8 years and 3.5 years respectively for the West African samples. Overall, there is statistically significant variation between countries in the wife's age at marriage according to an analysis of variance (p=.000). However, a larger proportion (83.2%) of the total variation occurs within rather than between countries.

4.7.8 Within-country variations in the age at marriage of wives of monogamous men

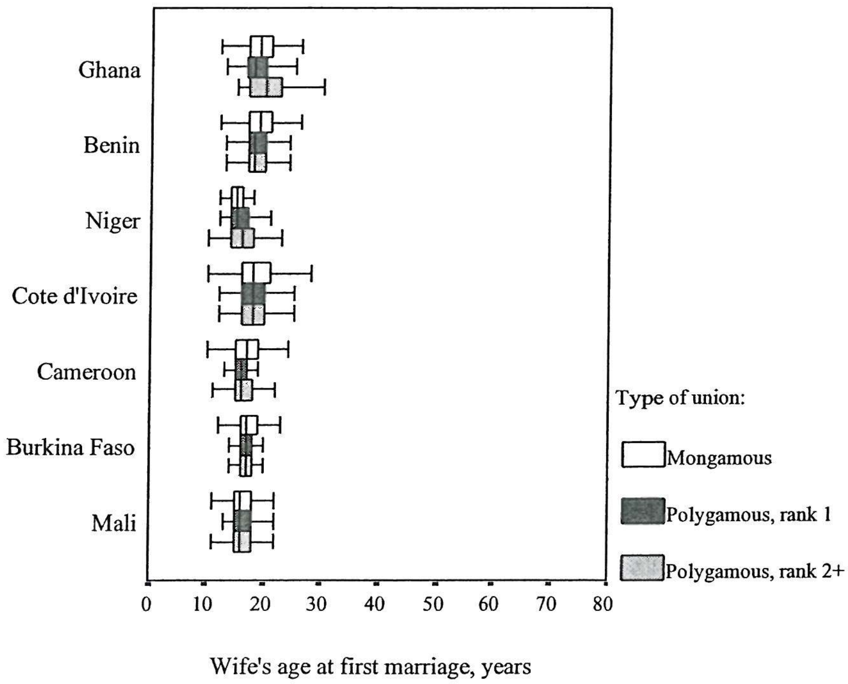
The amount of within-country variation in the age at marriage of wives of monogamous men is virtually the same as that observed for the age at marriage of all wives. Table 4.7 shows that, among the West African samples, where polygamy is most prevalent, there is at most 0.3 years difference in the magnitude of the IQR for all wives in comparison to just women married to monogamous husbands (for Burkina Faso and Niger). However, there is no consistency in the direction of this difference, i.e. for some samples, the IQR is slightly larger for the samples of monogamous women, while for some samples, the IQR is slightly smaller for the samples of monogamous women. In general, the observations made for the sample of all women apply to the sample of monogamous women. Again, there is statistically significant variation between countries in the age at marriage of women with monogamous husbands according to an analysis of variance (p=.000). As with the sample of all women, a larger proportion (83.3%) of the total variation occurs within rather than between countries. Figure 4.7 presents boxplots to show within-country variations in the age at marriage of wives for wives by union type, for countries in sub-Saharan Africa. As described, Figure 4.7 shows that there is seldom much difference in either the median or the distribution of the wife's age at marriage, for these three union types.

Figure 4.7 Boxplots to show the within-country variation in the wife's age at first marriage by union type for countries in sub-Saharan Africa

Central, East & South Africa



West Africa



Notes for Figure 4.7
See Appendix 4A for percentage distribution of women by union type and wife's rank
Kenya or Zimbabwe are not presented because these DHS surveys did not ask about the wife's rank
The same scale is used for the x-axis for comparison with the corresponding figure for the husband's age at marriage (Figure 4.6)

In sum, and as Casterline *et al* (1986) observed, there is greater variation within countries in the average age at marriage of husbands rather than wives. Conversely, there is greater cross-national variation in the average age at marriage of wives rather than husbands -an observation also made from the WISTAT3 analyses in Chapter 3. The next section considers the association between the age difference and the age at marriage of husbands and wives within countries.

4.8 The association between the age difference and age at marriage within countries

Casterline *et al* (1986) used the adjusted R-square statistic to quantify the proportion of the variation in the age difference within countries that can be accounted for by the age at marriage of husbands, and in turn wives. This analysis also makes use of this statistic from regressing the age difference on firstly, the husband's age at marriage, and then the wife's age at marriage. Table 4.8 gives the adjusted R-square for husbands and then wives for each country for samples of all couples and also samples of monogamous couples only.

As Casterline *et al* (1986) observed, a greater proportion of the variation in the age difference within countries can be accounted for by the husband's age at marriage than the wife's age at marriage. These proportions range from approximately one-quarter of the variation in the age difference for Kazakhstan (0.26) to almost all of the variation in the age difference in Burkina Faso (0.96). In contrast, the wife's age at marriage accounts for, at most, 15.3% of the variation in the age difference for the sample of couples in the Philippines. Indeed, for several samples, the wife's age at marriage accounts for virtually none of the variation in the age difference (Uganda, Zambia, Tanzania, and Ghana).

As discussed in Chapter 2, the regression of the age difference on, in turn the husband's age at marriage, and then the wife's age at marriage, can be simplified. When the age difference is regressed on the husband's age at marriage, this is statistically equivalent to the wife's age at marriage as the dependent variable because the husband's age at marriage appears on both sides of the equation. Since greater variation in the age at marriage of husbands than wives was observed within countries, it follows that the greater association within countries between the age difference and the husband's age at marriage reflects how the wife's age at marriage is easier to estimate with the husband's age at marriage, than *vice versa*.

Table 4.8 The proportion of the variation in the age difference that can be accounted for by the age at marriage of husbands and wives, for all couples and for monogamous couples only, by country

Region and country	Proportion of the variation in the age difference that can accounted for by the age at marriage of:			
	Husbands		Wives	
	All couples	Monogamous only	All couples	Monogamous only
Central & South America				
Guatemala	0.46	0.46	0.11	0.11
Peru	0.39	0.39	0.06	0.06
Brazil	0.49	0.49	0.07	0.07
Colombia	0.51	0.51	0.06	0.06
Dominican Republic	0.50	0.50	0.11	0.11
Haiti	0.63	0.63	0.09	0.09
West Asia & North Africa				
Kazakhstan	0.26	0.26	0.13	0.13
Uzbekistan	0.40	0.40	0.02	0.02
Egypt	0.51	0.51	0.06	0.06
South Asia				
Philippines	0.39	0.39	0.15	0.15
Nepal	0.61	0.58	0.02	0.02
Pakistan	0.61	0.57	0.02	0.02
Bangladesh	0.76	0.76	0.02	0.02
Central, East & South Africa				
Rwanda	0.71	0.66	0.11	0.13
Uganda	0.79	0.78	0.00	0.01
Malawi	0.68	0.64	0.02	0.02
Mozambique	0.77	0.70	0.02	0.01
Central African Republic	0.75	0.70	0.02	0.03
Zambia	0.65	0.62	0.00	0.01
Zimbabwe	0.79	0.65	0.02	0.02
Kenya	0.75	0.59	0.03	0.09
Tanzania	0.82	0.72	0.01	0.00
Comoros	0.73	0.73	0.02	0.01
West Africa				
Ghana	0.82	0.75	0.00	0.00
Benin	0.82	0.72	0.03	0.03
Niger	0.89	0.85	0.00	0.01
Burkina Faso	0.96	0.86	0.02	0.02
Cameroon	0.82	0.73	0.04	0.04
Côte d'Ivoire	0.79	0.64	0.04	0.04
Mali	0.87	0.78	0.01	0.02

Note for Table 4.8
The proportion of the variation that can be accounted for by the age at marriage of, in turn, husbands and wives is obtained from two separate linear regressions of the age difference on the age at marriage of husbands and then on the age at marriage of wives.

It is important to recall that the sample selection criteria used for this analysis and Casterline *et al*'s (1986) study only apply to women (see section 2.3.2). These criteria may limit the amount of variation in the age at marriage of wives while the age at marriage of husbands is not constrained in this way. This hypothesis can be examined with data from Kenya's DHS survey because this DHS survey included a male individual questionnaire and so the same sample selection criteria can be applied to husbands as well as wives.

There is highly statistically significant variation in the age at marriage of men who meet these sample selection criteria in comparison to those who did not (p=.000 according to a Mann-Whitney test). For men in their first marriage, the median age at marriage was 24.3 years (95% CI 24.0-24.6), while for men who had been married before, the median age at marriage was more than five years later at 29.4 years (95% CI 27.8-31.7). Furthermore, the variation in the age at marriage of husbands in their first marriage is much smaller than for men who had been married before. The IQR is 4.6 years (with lower and upper quartiles of 22.2 and 26.8, respectively) for men in their first marriage, while for men who had previously been married the IQR is 11.7 years (with lower and upper quartiles of 25.0 and 36.7, respectively).

Figure 4.8 shows the association between the age difference and the age at marriage of Kenyan husbands for those in their first marriage and then those who have previously been married. The corresponding adjusted R-square values are .46 and .73 respectively. Thus, the age at marriage of husbands accounts for approximately three-quarters of the variation in the age difference for couples where the husband has previously been married. In contrast, less than half of the variation in the age difference is accounted for by the age at marriage of husbands for couples where both the husband and wife are in their first marriage. When the age difference is regressed upon the age at marriage of wives, the corresponding adjusted R-square values are .01 and .19. Thus, the wife's age at marriage accounts for virtually none of the variation in the age difference when her husband has been married before. However, for couples where neither has previously been married, the wife's age at marriage accounts for almost a fifth of the variation in the age difference, although this is still smaller than the proportion accounted for by the husband's age at first marriage (0.46). To sum, the greater association between the age difference and the husband's age at marriage in comparison to the wife's age at marriage observed in Table 4.8 can only partly be attributed to the fact that sample selection criteria are applied only to wives.

Figure 4.8 Scatterplots showing the association between the age difference and the age at marriage of husbands by whether or not the husband was in his first marriage

Kenyan husbands in their first marriage (adjusted R-square: .46)

Kenyan husbands *not* in their first marriage (adjusted R-square: .73)

Note for Figure 4.8

Adjusted R-square statistics are obtained from the regression of the age difference on the age at marriage of husbands

This chapter has so far identified cross-national and within-country variations in the average age difference and the timing of marriage for men and women in 30 less developed countries. Given the age structure of the husbands and wives in these samples, the next section examines the hypothesis that the observed distributions of age differences have occurred randomly with respect to age.

4.9 Examining the hypothesis that union formation occurs randomly with respect to age

The hypothesis that the distribution of age differences occurs randomly with respect to age is examined by comparing the observed distribution of age differences that was discussed above, with the expected distribution of age differences. If this hypothesis is true then the observed and expected distributions will be equal. The expected distribution of age differences is obtained for each sample by crossing the female and male age at marriage distributions. This is the approach that Casterline *et al* (1986) used, which they expressed as

$$d'_k = \sum_{i-j} h_i w_j$$

Source: Casterline *et al* (1986, p. 362)

Where h_i is the proportion of husbands who married at age i where i has the same age range as the observed age at marriage distribution for husbands. Similarly, w_j is the proportion of wives who married at age j where j has the same age range as the observed age at marriage distribution for wives. Thus, i-j is the age difference calculated as the husband's age minus the wife's age. The summation is across values of i and j for which i-j=k.

In order to enable comparisons with the WFS study, Table 4.9 replicates a comparable table presented by Casterline *et al* (1986). It shows the median, lower and upper quartiles, the IQR and the percent of couples with a negative age difference and the percent of couples with an age difference of at least ten years for the observed and expected distributions of all couples.
In addition, and unlike Casterline *et al* (1986), Table 4.10 shows these summary statistics for the observed and expected distributions of age differences of just monogamous couples.

Table 4.9: Comparison of observed & expected age difference distributions: All couples

	Median difference (years)			Inter-quartile range (years)			Lower quartile (years)			Upper quartile (years)			% <0 years (years)			% 10+ years (years)		
	Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp
Central & South America																		
Guatemala	2.1	2.7	-0.6	5.0	7.4	-2.3	0.1	-0.9	1.0	5.1	6.5	-1.4	13.4	24.4	-11.0	9.4	14.6	-5.2
Peru	2.8	2.4	0.4	5.6	9.1	-3.5	0.0	-1.6	1.6	5.6	7.5	-1.9	14.8	27.9	-13.1	11.2	19.5	-8.3
Brazil	2.9	3.2	-0.3	6.3	9.4	-3.1	0.1	-1.3	1.5	6.4	8.1	-1.7	16.7	26.4	-9.7	13.6	21.5	-7.9
Colombia	3.1	3.6	-0.5	6.4	9.5	-3.1	0.3	-0.9	1.2	6.7	8.5	-1.8	13.3	24.6	-11.3	16.9	23.1	-6.2
Dominican Republic	3.9	4.6	-0.7	6.4	9.3	-2.8	1.3	0.2	1.1	7.7	9.5	-1.7	7.9	19.6	-11.7	20.5	26.9	-6.4
Haiti	4.1	4.5	-0.4	6.4	9.5	-3.2	1.1	-0.1	1.2	7.5	9.5	-2.0	10.9	20.9	-10.0	18.2	26.8	-8.6
Regional	*3.2*	*3.5*	*-0.3*	*6.0*	*9.0*	*-3.0*	*0.5*	*-0.8*	*1.3*	*6.5*	*8.3*	*-1.8*	*12.8*	*24.0*	*-11.1*	*15.0*	*22.1*	*-7.1*
West Asia & North Africa																		
Kazakhstan	1.8	2.1	-0.3	3.8	5.7	-1.9	0.0	-0.7	0.7	3.8	5.0	-1.2	10.9	23.0	-12.1	3.6	7.4	-3.8
Uzbekistan	2.3	2.5	-0.2	3.2	4.7	-1.5	0.7	0.1	0.6	3.9	4.8	-0.9	2.4	16.3	-13.9	2.6	5.8	-3.2
Egypt	5.7	6.1	-0.4	6.0	8.5	-2.5	2.9	2.9	0.0	8.9	10.4	-1.5	2.6	13.3	-10.7	24.1	32.3	-8.2
Regional	*3.3*	*3.6*	*-0.3*	*4.3*	*6.3*	*-2.0*	*1.2*	*0.8*	*0.4*	*5.5*	*6.7*	*-1.2*	*5.3*	*17.5*	*-12.2*	*10.1*	*15.2*	*-5.1*
South Asia																		
Philippines	2.1	1.6	0.4	5.3	8.4	-3.1	-0.2	-2.5	2.3	5.1	5.9	-0.8	17.5	33.1	-15.6	9.3	12.6	-3.3
Nepal	2.9	3.1	-0.2	4.3	6.2	-1.9	1.0	0.1	0.9	5.3	6.3	-1.0	6.9	17.5	-10.6	8.6	13.0	-4.4
Pakistan	4.4	4.9	-0.5	4.8	8.6	-3.8	2.9	0.7	2.2	7.7	9.3	-1.6	5.4	17.2	-11.8	19.7	26.4	-6.7
Bangladesh	7.7	8.0	-0.3	5.3	6.9	-1.6	5.4	4.7	0.6	10.7	11.7	-1.0	0.1	3.2	-3.1	36.7	42.0	-5.3
Regional	*4.3*	*4.4*	*-0.1*	*4.9*	*7.5*	*-2.6*	*2.3*	*0.8*	*1.5*	*7.2*	*8.3*	*-1.1*	*7.5*	*17.8*	*-10.3*	*18.6*	*23.5*	*-4.9*
Central, East & South Africa																		
Rwanda	3.8	4.0	-0.2	6.3	7.6	-1.3	0.8	0.4	0.4	7.1	7.9	-0.9	11.8	17.4	-5.6	16.3	20.0	-3.7
Uganda	4.3	4.6	-0.3	5.4	7.3	-2.0	2.0	1.2	0.8	7.4	8.6	-1.2	3.7	12.9	-9.2	17.3	22.9	-5.6
Malawi	4.5	4.8	-0.3	5.3	6.1	-0.8	2.5	2.0	0.5	7.7	8.0	-0.3	1.1	9.1	-8.0	16.5	22.6	-6.1
Mozambique	4.8	5.4	-0.6	6.9	8.7	-1.8	2.2	1.6	0.6	9.1	10.3	-1.1	3.4	12.5	-9.1	25.7	30.0	-4.3
Central African Republic	4.9	5.4	-0.5	7.0	8.6	-1.6	2.2	1.5	0.7	9.2	10.2	-1.0	3.9	12.9	-9.0	25.8	29.9	-4.1
Zambia	4.9	5.2	-0.3	4.6	6.9	-2.3	3.0	2.0	1.1	7.6	8.9	-1.3	0.8	10.1	-9.3	17.3	24.4	-7.1
Zimbabwe	5.4	5.9	-0.5	6.2	8.3	-2.1	2.8	2.0	0.7	8.9	10.3	-1.4	2.5	10.9	-8.4	24.7	31.0	-6.3
Kenya	5.4	5.9	-0.5	5.7	7.6	-1.9	3.1	2.3	0.8	8.8	9.9	-1.1	2.0	9.9	-7.9	23.8	29.7	-5.9
Tanzania	5.9	6.3	-0.4	6.9	8.6	-1.7	3.0	2.4	0.6	9.9	11.0	-1.1	1.6	10.0	-8.4	30.0	33.8	-3.8
Comores	7.8	8.3	-0.5	9.3	11.9	-2.7	4.0	2.9	1.0	13.2	14.8	-1.6	3.6	11.9	-8.3	43.0	46.9	-3.9
Regional	*5.2*	*5.6*	*-0.4*	*6.3*	*8.1*	*-1.8*	*2.6*	*1.8*	*0.7*	*8.9*	*10.0*	*-1.1*	*3.4*	*11.8*	*-8.3*	*24.0*	*29.1*	*-5.1*
West Africa																		
Ghana	5.9	6.4	-0.5	6.8	9.1	-2.3	3.2	2.3	0.9	10.0	11.4	-1.4	1.4	10.4	-9.0	30.2	35.6	-5.4
Benin	6.4	7.0	-0.7	7.3	9.5	-2.1	3.5	2.8	0.7	10.8	12.3	-1.4	2.0	9.4	-7.4	33.8	39.1	-5.3
Niger	8.6	8.7	-0.2	8.8	8.7	0.1	4.1	4.9	-0.8	12.8	13.6	-0.7	1.5	4.2	-2.7	46.8	48.0	-1.2
Burkina Faso	8.7	9.3	-0.5	11.2	11.6	-0.4	5.0	5.0	0.0	16.3	16.6	-0.3	0.5	2.8	-2.3	48.4	51.3	-2.9
Cameroon	9.1	9.8	-0.6	9.6	10.3	-0.7	4.8	4.9	-0.1	14.3	15.2	-0.8	0.6	6.4	-5.8	51.4	54.2	-2.8
Cote d'Ivoire	9.2	9.8	-0.6	9.8	11.3	-1.5	5.3	4.7	0.6	15.1	15.9	-0.9	2.3	7.3	-5.0	50.9	53.8	-2.9
Mali	9.7	10.1	-0.4	8.2	9.8	-1.6	6.5	6.1	0.4	14.6	15.8	-1.2	0.2	3.0	-2.8	56.1	56.7	-0.6
Regional	*8.2*	*8.7*	*-0.5*	*8.8*	*10.0*	*-1.2*	*4.6*	*4.4*	*0.2*	*13.4*	*14.4*	*-1.0*	*1.2*	*6.2*	*-5.0*	*45.4*	*48.4*	*-3.0*

Note for Table 4.9

Countries are presented in increasing order of observed age difference width, by region.

Regions are defined according to the United Nations Statistics Division

Table 4.10: Comparison of observed & expected age difference distributions: Monogamous couples only

	Median difference (years)			Inter-quartile range (years)			Lower quartile (years)			Upper quartile (years)			% <0 (years)			% 10+ (years)		
	Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp
Central & South America																		
Guatemala	2.1	2.7	-0.6	5.0	7.4	-2.3	0.1	-0.9	1.0	5.1	6.5	-1.4	13.4	24.4	-11.0	9.4	14.6	-5.2
Peru	2.8	2.4	0.4	5.6	9.1	-3.5	0.0	-1.6	1.6	5.6	7.5	-1.9	14.8	27.9	-13.1	11.2	19.5	-8.3
Brazil	2.9	3.2	-0.3	6.3	9.4	-3.1	0.1	-1.3	1.5	6.4	8.1	-1.7	16.7	26.4	-9.7	13.6	21.5	-7.9
Colombia	3.1	3.6	-0.5	6.4	9.5	-3.1	0.3	-0.9	1.2	6.7	8.5	-1.8	13.3	24.6	-11.3	16.9	23.1	-6.2
Dominican Republic	3.9	4.6	-0.7	6.4	9.3	-2.8	1.3	0.2	1.1	7.7	9.5	-1.7	7.9	19.6	-11.7	20.5	26.9	-6.4
Haiti	4.1	4.5	-0.4	6.4	9.5	-3.2	1.1	-0.1	1.2	7.5	9.5	-2.0	10.9	20.9	-10.0	18.2	26.8	-8.6
Regional	*3.2*	*3.5*	*-0.3*	*6.0*	*9.0*	*-3.0*	*0.5*	*-0.8*	*1.3*	*6.5*	*8.3*	*-1.8*	*12.8*	*24.0*	*-11.1*	*15.0*	*22.1*	*-7.1*
West Asia & North Africa																		
Kazakhstan	1.8	2.1	-0.3	3.8	5.7	-1.9	0.0	-0.7	0.7	3.8	5.0	-1.2	10.9	23.0	-12.1	3.6	7.4	-3.8
Uzbekistan	2.3	2.5	-0.2	3.2	4.7	-1.5	0.7	0.1	0.6	3.9	4.8	-0.9	2.4	16.3	-13.9	2.6	5.8	-3.2
Egypt	5.7	6.1	-0.4	6.0	8.5	-2.5	2.9	2.9	0.0	8.9	10.4	-1.5	2.6	13.3	-10.7	24.1	32.3	-8.2
Regional	*3.3*	*3.6*	*-0.3*	*4.3*	*6.3*	*-2.0*	*1.2*	*0.8*	*0.4*	*5.5*	*6.7*	*-1.2*	*5.3*	*17.5*	*-12.2*	*10.1*	*15.2*	*-5.1*
South Asia																		
Philippines	2.1	1.6	0.4	5.3	8.4	-3.1	-0.2	-2.5	2.2	5.1	5.9	-0.8	17.5	33.1	-15.6	9.3	12.6	-3.3
Nepal	2.8	3.0	-0.2	4.1	6.0	-1.8	1.0	0.1	0.9	5.2	6.1	-1.0	6.9	17.5	-10.6	7.5	11.9	-4.4
Pakistan	4.3	4.7	-0.4	4.6	8.4	-3.8	2.8	0.6	2.3	7.4	9.0	-1.6	5.6	17.5	-11.9	18.0	25.0	-7.0
Bangladesh	7.7	8.0	-0.3	5.3	6.9	-1.6	5.4	4.7	0.6	10.7	11.7	-1.0	0.1	3.2	-3.1	36.7	42.0	-5.3
Regional	*4.2*	*4.3*	*-0.1*	*4.8*	*7.4*	*-2.6*	*2.3*	*0.7*	*1.5*	*7.1*	*8.2*	*-1.1*	*7.5*	*17.8*	*-10.3*	*17.9*	*22.9*	*-5.0*
Central, East & South Africa																		
Rwanda	3.6	3.8	-0.2	5.9	7.3	-1.4	0.7	0.3	0.4	6.6	7.6	-1.0	12.1	17.7	-5.6	14.3	18.3	-4.0
Uganda	4.0	4.3	-0.3	4.8	6.7	-1.9	1.9	1.2	0.7	6.7	7.8	-1.1	3.8	12.7	-8.9	14.0	19.2	-5.2
Malawi	4.4	4.7	-0.3	5.0	6.2	-1.2	2.3	1.8	0.5	7.3	8.1	-0.7	1.3	9.3	-8.0	14.4	20.3	-5.9
Mozambique	4.1	4.6	-0.5	6.0	8.0	-1.9	2.0	1.1	0.9	8.0	9.1	-1.1	4.0	14.2	-10.2	20.8	25.3	-4.5
Central African Republic	4.5	5.0	-0.5	6.0	7.9	-1.9	2.0	1.3	0.7	8.0	9.2	-1.2	4.3	13.5	-9.2	20.8	25.8	-5.0
Zambia	4.9	5.2	-0.3	4.5	6.7	-2.3	3.0	2.0	1.0	7.4	8.7	-1.3	0.8	10.0	-9.2	16.1	23.4	-7.3
Zimbabwe	5.0	5.3	-0.3	5.4	7.5	-2.2	2.6	1.7	0.9	8.0	9.2	-1.3	2.7	11.7	-9.0	19.7	26.2	-6.5
Kenya	5.0	5.5	-0.5	5.0	6.9	-1.9	2.9	2.1	0.8	7.9	9.0	-1.1	2.1	10.0	-7.9	18.6	25.2	-6.6
Tanzania	5.3	5.6	-0.3	6.0	7.6	-1.6	2.7	2.0	0.7	8.8	9.7	-0.9	1.7	10.6	-8.9	23.4	28.3	-4.9
Comores	7.1	7.7	-0.6	8.3	10.9	-2.6	3.6	2.6	1.0	11.9	13.5	-1.7	3.6	11.9	-8.3	40.0	43.6	-3.6
Regional	*4.8*	*5.2*	*-0.4*	*5.7*	*7.6*	*-1.9*	*2.4*	*1.6*	*0.8*	*8.1*	*9.2*	*-1.1*	*3.6*	*12.2*	*-8.5*	*20.2*	*25.6*	*-5.4*
West Africa																		
Ghana	5.5	5.9	-0.4	6.1	8.4	-2.3	2.9	2.0	0.9	9.0	10.4	-1.4	1.6	11.2	-9.6	25.0	31.6	-6.6
Benin	5.2	5.8	-0.7	5.8	8.2	-2.5	3.0	2.0	1.0	8.8	10.2	-1.5	1.9	11.5	-9.6	24.3	31.0	-6.7
Niger	7.7	7.8	-0.1	7.2	7.2	0.0	3.7	4.4	-0.8	10.9	11.6	-0.7	1.8	4.3	-2.5	39.4	41.2	-1.8
Burkina Faso	6.4	6.7	-0.3	5.6	6.9	-1.3	4.0	3.5	0.4	9.5	10.4	-0.9	0.3	4.5	-4.2	27.9	33.8	-5.9
Cameroon	7.2	7.9	-0.7	6.7	8.4	-1.7	4.1	3.7	0.4	10.8	12.1	-1.3	0.6	8.5	-7.9	39.2	43.2	-4.0
Cote d'Ivoire	7.5	8.0	-0.5	7.2	9.6	-2.4	4.5	3.4	1.1	11.6	13.0	-1.4	2.9	9.7	-6.8	39.9	44.4	-4.5
Mali	8.5	8.7	-0.3	7.2	7.6	-0.4	5.8	5.2	0.5	13.0	12.9	0.1	0.2	3.6	-3.4	46.5	47.9	-1.4
Regional	*6.9*	*7.3*	*-0.4*	*6.5*	*8.0*	*-1.5*	*4.0*	*3.5*	*0.5*	*10.5*	*11.5*	*-1.0*	*1.3*	*7.6*	*-6.3*	*34.6*	*39.0*	*-4.4*

Note for Table 4.10

Countries are presented in increasing order of observed age difference width among all couples, by region for comparison with Table 4.9

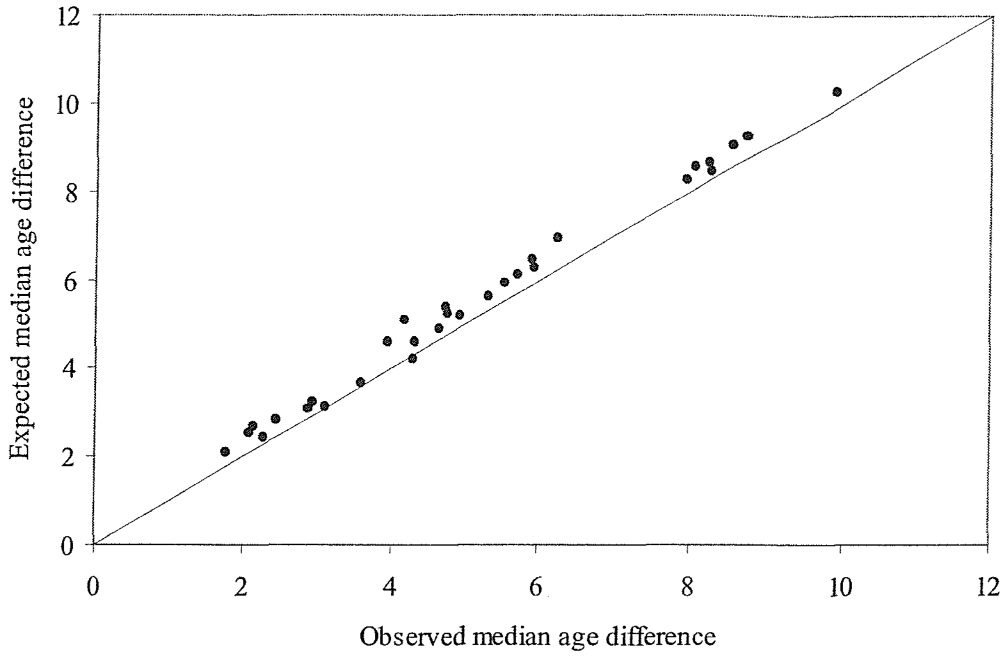
Regions are defined according to the United Nations Statistics Division

4.9.1 Cross-national variations in the observed and expected age difference distributions

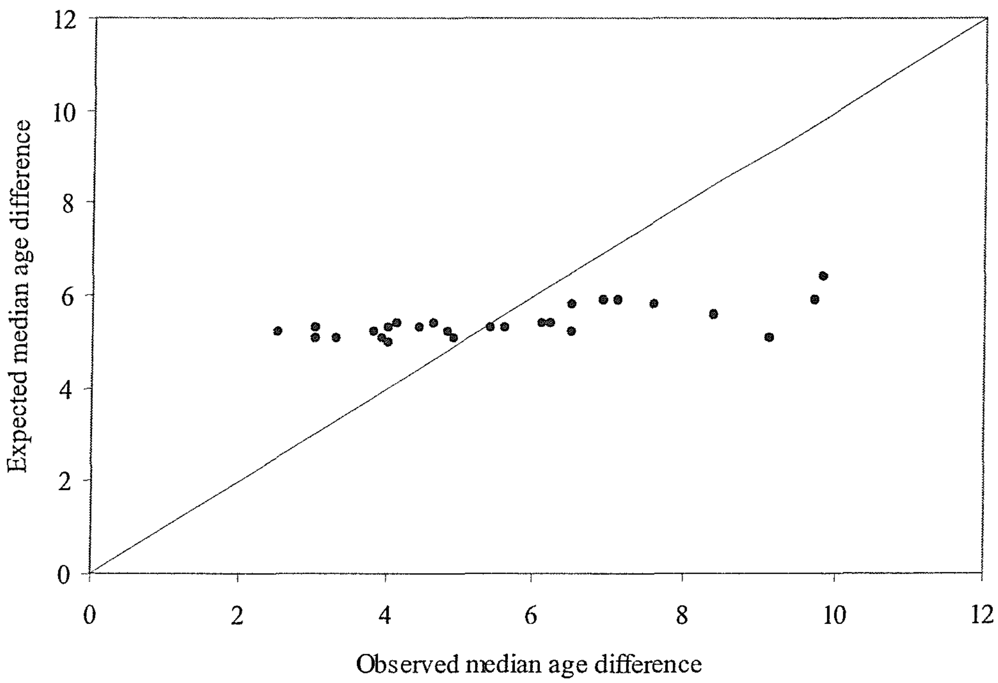
Tables 4.9 and 4.10 show that the observed median age difference is less than 0.5 years smaller than the expected median age difference for almost all countries, regardless of whether the sample consists of all couples or just monogamous couples. Thus, in general, the observed and expected median age differences are very closely related. In contrast, Casterline *et al* (1986) reported that the expected medians were quite homogeneous for the 28 countries that they studied. Their expected medians ranged from 5.0 years to 6.4 years while their observed medians ranged from 2.5 years to 9.8 years. Figure 4.9 shows plots of the observed median and expected median for the two studies to illustrate this discrepancy. Given this discrepancy, it should be stressed that the estimates obtained from the analysis of the DHS data were checked a number of times by the author and also independently.

Figure 4.9 Observed median age difference by expected median age difference

Estimates for all couples from 30 DHS surveys



Estimates for all couples from 28 WFS surveys



Notes for Figure 4.9
WFS estimates as reported by Casterline *et al* (1986, p.363)
This method of graphical display was originally presented in Ní Bhrolcháin (1997)

4.9.2 Re-analysis of the WFS survey data

Given this discrepancy between the two studies, the observed and expected age difference distributions were calculated using WFS data for five of the 28 WFS surveys that Casterline *et al* (1986) studied. These are Bangladesh, Colombia, Egypt, Ghana and the Philippines. The same sample selection criteria were used (see section 2.3.2) to ensure consistency with the methods reported by Casterline *et al* (1986). However, as Table 4.11 shows, there is a slight discrepancy between the weighted number of cases used in this re-analysis and as reported by Casterline *et al* (1986, p. 357).

Table 4.11 gives the summary statistics for this re-analysis as well as the results reported by Casterline *et al* (1986) for comparison. Of the five countries that were re-analysed, the expected medians range from 2.4 years for the Philippines to 8.8 years for Bangladesh. However, the expected median age difference for these five countries as reported by Casterline *et al* (1986) range from 5.1 years for Bangladesh to 5.8 years for Ghana. Thus, there is as much as 3.7 years difference (for Bangladesh) between the expected medians that Casterline *et al* (1986) report and those obtained from re-running their analyses. Figure 4.10 plots the observed and expected median age differences from these two sources to illustrate this point.

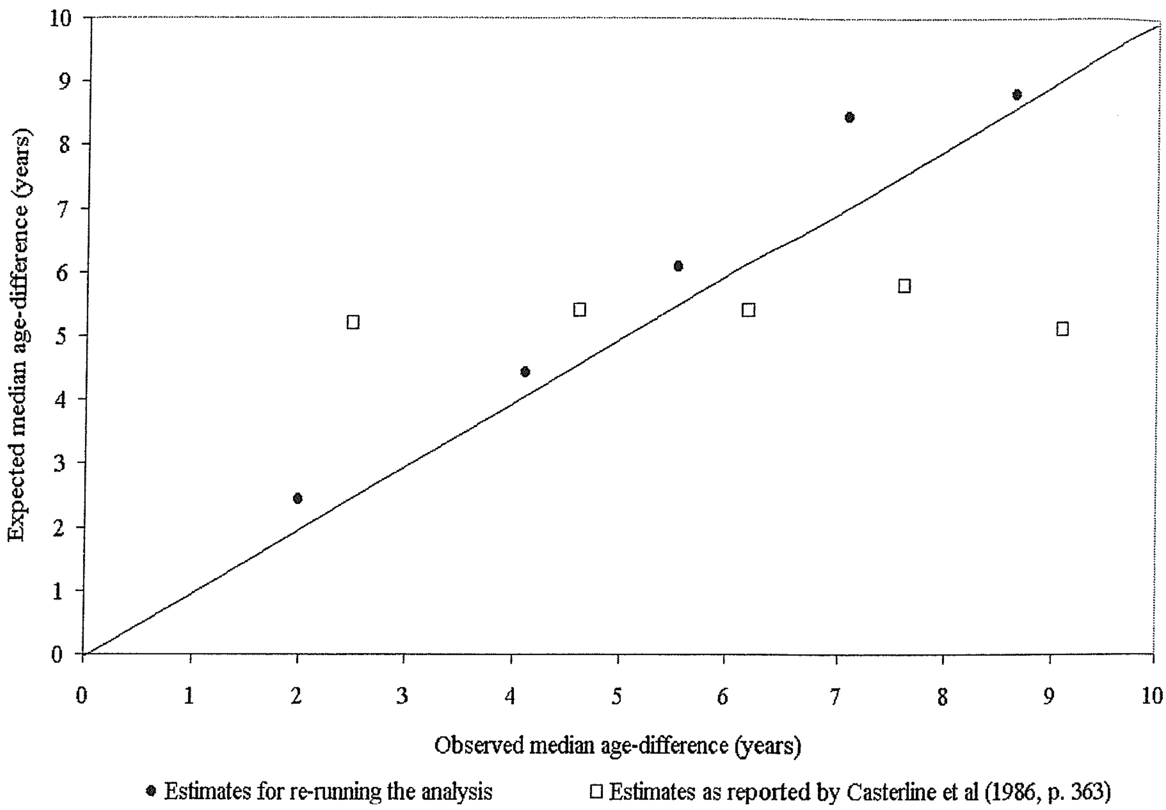
It is assumed that Casterline *et al* (1986) reviewed their methods and results, especially as they described the uniformity of their expected medians as *"the most remarkable feature... and wholly unexpected"* (Casterline *et al*, 1986, p. 362). However, based on these findings they rejected the hypothesis that variations in the age difference reflect the age-sex structure of those marrying. Importantly, the same conclusion can not be made from this chapter's analysis of the DHS data and the re-analysis of the data from five WFS surveys.

The next section considers whether the observed variations in the distributions of age differences are as expected.

Table 4.11: Comparison of observed & expected age difference distributions

World Fertility Survey	Sample size (weighted)	Median difference (years)			Inter-quartile range (years)			Lower quartile (years)			Upper quartile (years)			% <0 years (years)			% 10+ years (years)		
		Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp	Obs	Exp	Obs-Exp
Bangladesh																			
2001 re-run	2029	8.6	8.8	-0.2	5.6	7.1	-1.5	6.1	5.5	0.5	11.7	12.6	-1.0	0.0	2.1	-2.1	46.6	48.5	-1.9
Casterline *et al*	*1987*	9.1	5.1	4.0	5.6	8.8	-3.2	Not available			Not available			0.2	19.0	-18.8	47.0	25.0	22.0
N difference	*42*	*-0.5*	*3.7*	*-4.2*	*0.0*	*-1.7*	*1.7*							*-0.2*	*-16.9*	*16.7*	*-0.4*	*23.5*	*-23.9*
% difference	*2.1%*																		
Colombia																			
2001 re-run	1208	4.1	4.4	-0.3	6.7	9.0	-2.3	1.2	0.2	1.0	7.9	9.2	-1.3	10.3	19.4	-9.1	20.5	25.7	-5.2
Casterline *et al*	*1342*	4.6	5.4	-0.8	6.6	9.6	-3.0	Not available			Not available			10.0	20.0	-10.0	20.0	28.0	-8.0
N difference	*-134*	*-0.5*	*-1.0*	*0.5*	*0.1*	*-0.6*	*0.7*							*0.3*	*-0.6*	*0.9*	*0.5*	*-2.3*	*2.8*
% difference	*10.0%*																		
Egypt																			
2001 re-run	2919	5.5	6.1	-0.6	6.3	8.9	-2.5	2.9	1.8	1.1	9.2	10.6	-1.4	2.5	13.2	-10.7	26.4	32.9	-6.5
Casterline *et al*	*3332*	6.2	5.4	0.8	6.4	9.2	-2.8	Not available			Not available			2.4	21.0	-18.6	27.0	19.0	8.0
N difference	*-413*	*-0.7*	*0.7*	*-1.4*	*-0.1*	*-0.3*	*0.3*							*0.1*	*-7.8*	*7.9*	*-0.6*	*13.9*	*-14.5*
% difference	*12.4%*																		
Ghana																			
2001 re-run	1629	7.1	8.5	-1.4	8.1	9.5	-1.4	4.2	4.3	-0.1	12.3	13.8	-1.5	0.9	5.4	-4.5	39.7	46.8	-7.1
Casterline *et al*	*1779*	7.6	5.8	1.8	8.0	11.1	-3.1	Not available			Not available			1.0	21.0	-20.0	39.0	33.0	6.0
N difference	*-150*	*-0.5*	*2.7*	*-3.2*	*0.1*	*-1.6*	*1.7*							*-0.1*	*-15.6*	*15.5*	*0.7*	*13.8*	*-13.1*
% difference	*8.4%*																		
Philippines																			
2001 re-run	3508	2.0	2.4	-0.4	5.3	8.5	-3.2	-0.2	-1.7	1.5	5.0	6.8	-1.7	16.8	28.5	-11.7	9.3	16.5	-7.2
Casterline *et al*	*3448*	2.5	5.2	-2.7	5.2	8.7	-3.5	Not available			Not available			15.9	19.0	-3.1	10.0	26.0	-16.0
N difference	60	*-0.5*	*-2.8*	*2.3*	*0.1*	*-0.2*	*0.3*							*0.9*	*9.5*	*-8.6*	*-0.7*	*-9.5*	*8.8*
% difference	*1.7%*																		

Figure 4.10 Observed median age difference by expected median age difference: Comparing estimates for five of the 28 WFS surveys reported by Casterline *et al* (1986, p.363) and from re-running their analysis



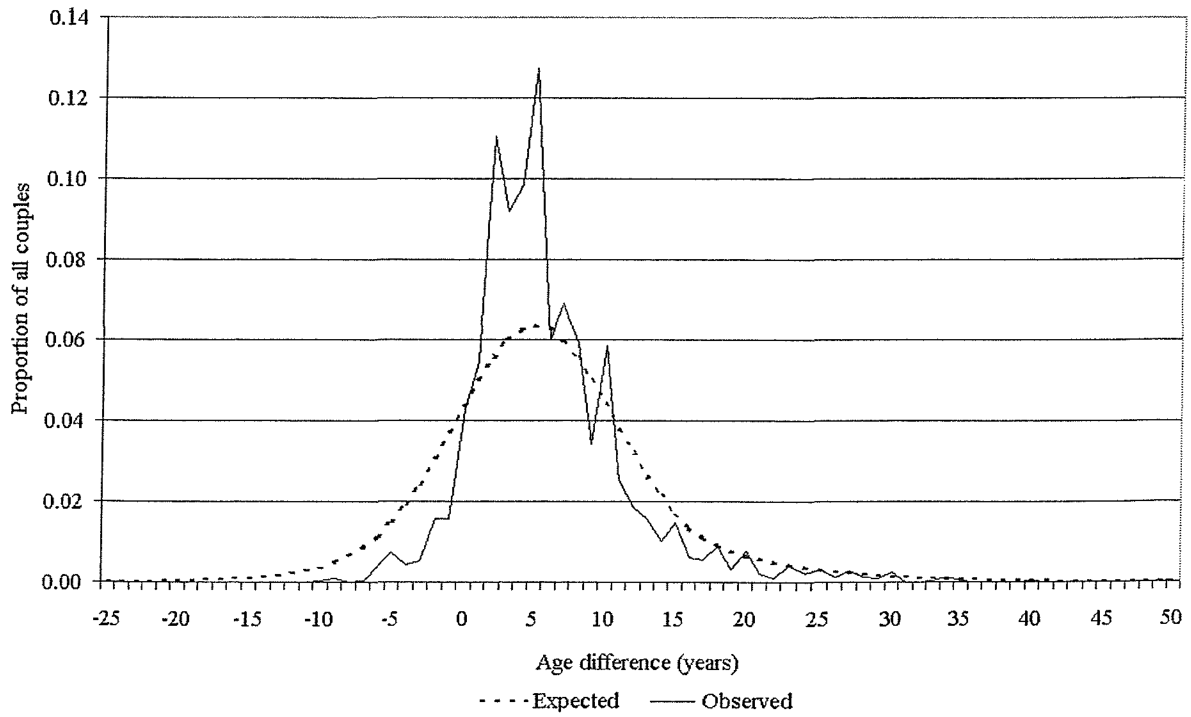
4.9.3 Within-country variations in the observed and expected age difference distributions

The inter-quartile range (IQR) is used as a measure of variation in the observed and expected distributions of the age difference for these 30 countries. Table 4.9 shows that the observed IQR is smaller than expected for all but one of the samples studied. The magnitude of this difference ranges from 0.4 years for Burkina Faso to 3.8 years for Pakistan. Thus, for example, in the sample of couples in Pakistan, it is expected that half of these couples will have an age difference of between 0.7 years and 9.3 years, but instead it is observed that half of the couples have an age difference of between 2.9 years and 7.7 years. This pattern exists for almost all samples, regardless of whether the sample consists of all couples or just monogamous couples. An exception is Niger, where the observed IQR for all couples is larger than expected albeit by 0.1 years relative to an expected IQR of 8.7 years. Among monogamous couples in Niger, the observed and expected IQRs are both estimated as 7.2 years.

Despite the inconsistency with Casterline *et al*'s (1986) results regarding the expected medians, they also reported that the observed IQRs were smaller than expected for the 28 samples that they studied. The difference between the observed and expected IQRs for their 28 samples ranged from 1.8 years for Korea to 4.7 years for Haiti. Their expected IQR estimates are more heterogeneous than their expected medians.

Table 4.9 shows that for most samples the smaller than expected IQR results from a larger than expected lower quartile and a smaller than expected upper quartile. Among the samples of all couples, Niger and Cameroon are the only samples where the observed lower quartile is smaller than expected at 0.8 and 0.1 years, respectively. Among the samples of monogamous couples, Niger is the only sample where the observed lower quartile is smaller than expected, also by 0.8 years. Generally the observed age difference distribution is more concentrated around the observed median age difference than expected. Figure 4.11 shows the difference in the observed and expected distributions for Pakistan, the sample with the greatest difference between the observed and expected IQRs.

Figure 4.11 Line graph showing the observed and expected age difference distributions: Pakistan



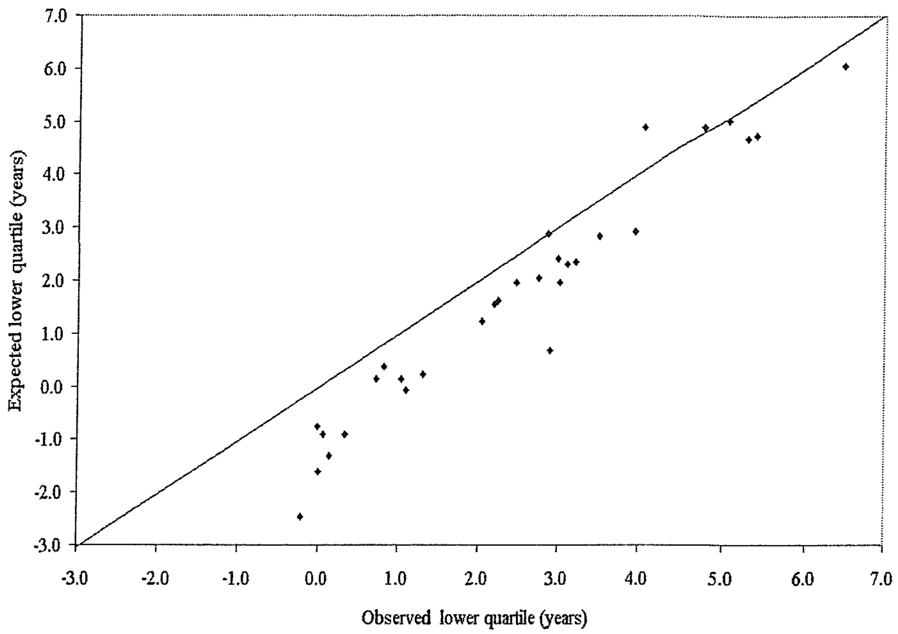
There is little association between the magnitude of the difference between the observed and expected lower quartiles and the median observed age difference. The sample with the greatest difference between the observed and expected lower quartiles is the Philippines where the lower quartile is observed as -0.2 years, which compares to an expected lower quartile of –2.5 years. This means that a quarter of couples in the Philippines sample are expected to have an age difference where the wife is at least 2.5 years older than her husband, but instead, it is observed that approximately one in ten couples have such age differences. While the median observed age difference for the Philippines sample is relatively small at 2.1 years, other samples with medians of a similar magnitude such as Guatemala and Uzbekistan, do not have such a great difference between the observed and expected lower quartiles, with differences of 1.0 years and 0.6 years respectively. Among the West African samples where median age differences are relatively large, there is some variation in the difference between the observed and expected lower quartiles, ranging from a difference of -0.8 years for Niger to +0.9 years for Ghana. These patterns also apply when focussing on monogamous couples.

For almost all samples of all couples and those of just monogamous couples, the upper quartile in the observed distribution of the age difference is smaller than expected. The greatest difference between the observed and expected upper quartiles is identified among the sample of couples from Haiti where the observed upper quartile, 7.5 years, is 2.0 years smaller than expected. This means that three-quarters of couples are expected to have an age difference of at most, 9.5 years, but instead it is observed that three-quarters of couples have an age difference of at most, 7.5 years. Approximately 84% of couples have an age difference of at most 9.5 years.

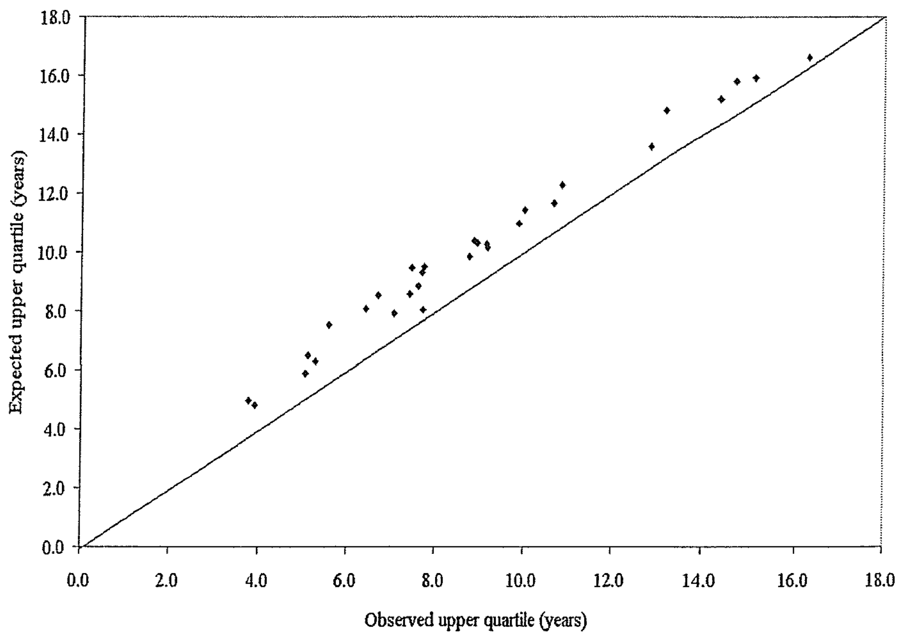
The smallest difference between the observed and expected upper quartiles is identified for two samples of all couples: Malawi and Burkina Faso. The observed upper quartiles of, respectively, 7.7 years and 16.3 years, are both 0.3 years smaller than expected. The only sample where the observed upper quartile is larger than expected is the sample of monogamous couples in Mali albeit by 0.1 years. Figure 4.12 illustrates these results with a scatterplot of the observed lower quartiles plotted against the expected lower quartiles, and the observed upper quartiles plotted against the expected upper quartiles, for all 30 samples. Unfortunately, comparisons can not be made with Casterline *et al*'s (1986) study in terms of the expected lower and upper quartiles as they do not report these statistics.

Figure 4.12 Scatterplots showing the association between the observed and expected quartiles from the observed and expected distributions of age differences for all couples in samples from 30 countries

Lower quartile

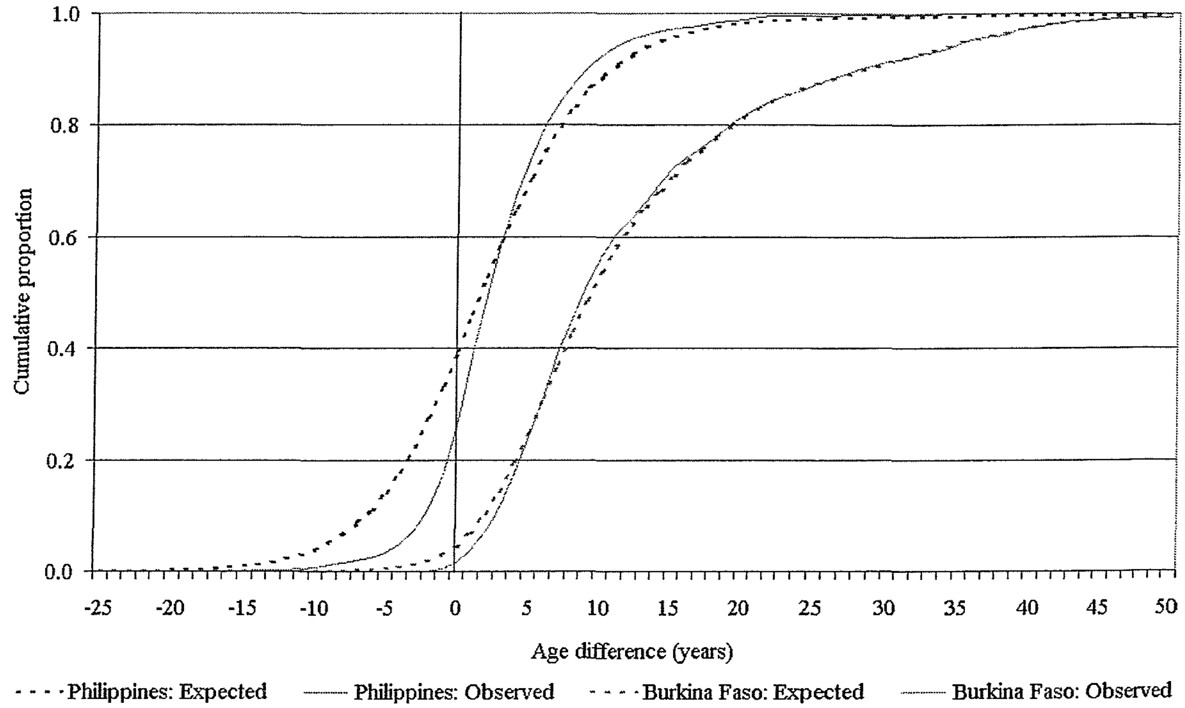


Upper quartile



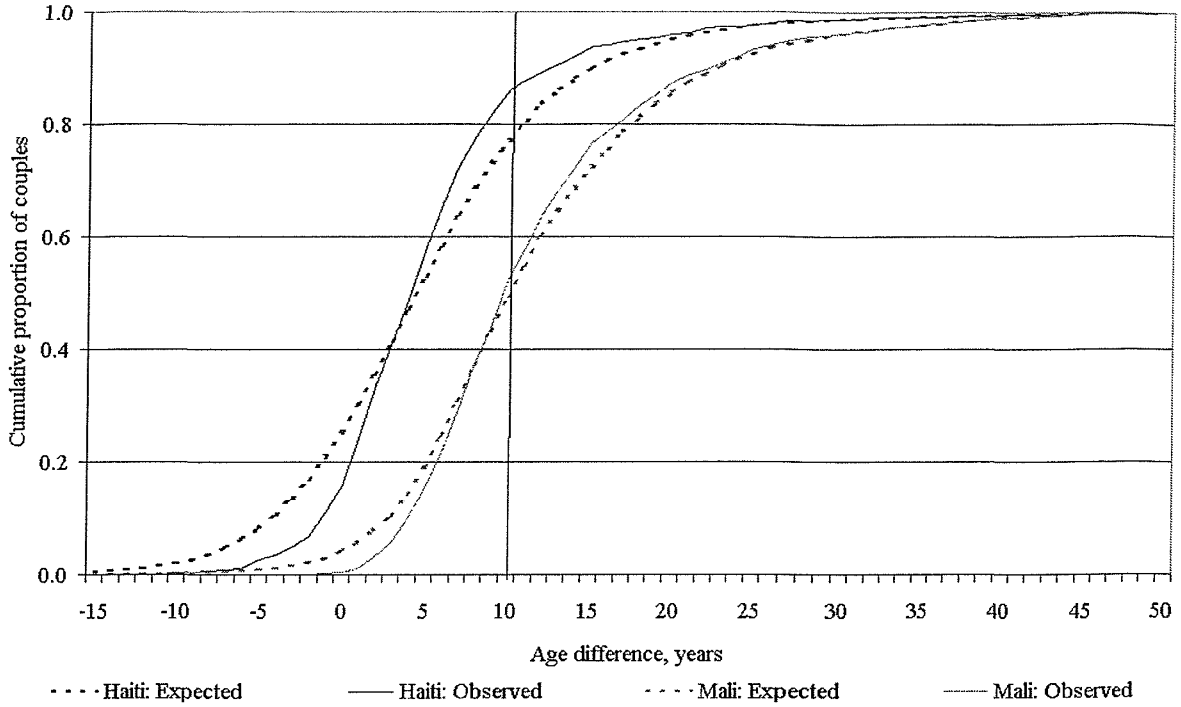
Casterline *et al* (1986) also considered differences between the observed and expected distributions in terms of more extreme age differences, such as negative age differences and those of at least ten years. The observed proportion is smaller than expected for both these extreme age differences for all 30 DHS surveys studied. This difference between the observed and expected proportions with a negative age difference ranges from 2.3% for Burkina Faso to 15.6% for the Philippines, and is evident from Figure 4.13.

Figure 4.13 The observed and expected cumulative distributions of the age difference: The proportion of couples with a negative age difference in the Philippines sample in comparison to the Burkina Faso sample



The difference between the observed and expected proportions of age differences of at least ten years ranges from 0.6% for Mali to 8.6% for Haiti (Figure 4.14). For almost all samples, there is a greater difference between the observed and expected proportions of negative age differences than the observed and expected proportions of age differences of at least ten years. Burkina Faso and Bangladesh are the only two countries where the converse applies. Thus in these two countries, the expected proportion of couples with a negative age difference is closer to the observed proportion than the expected proportion of couples with an age difference of at least ten years is to the corresponding observed figure.

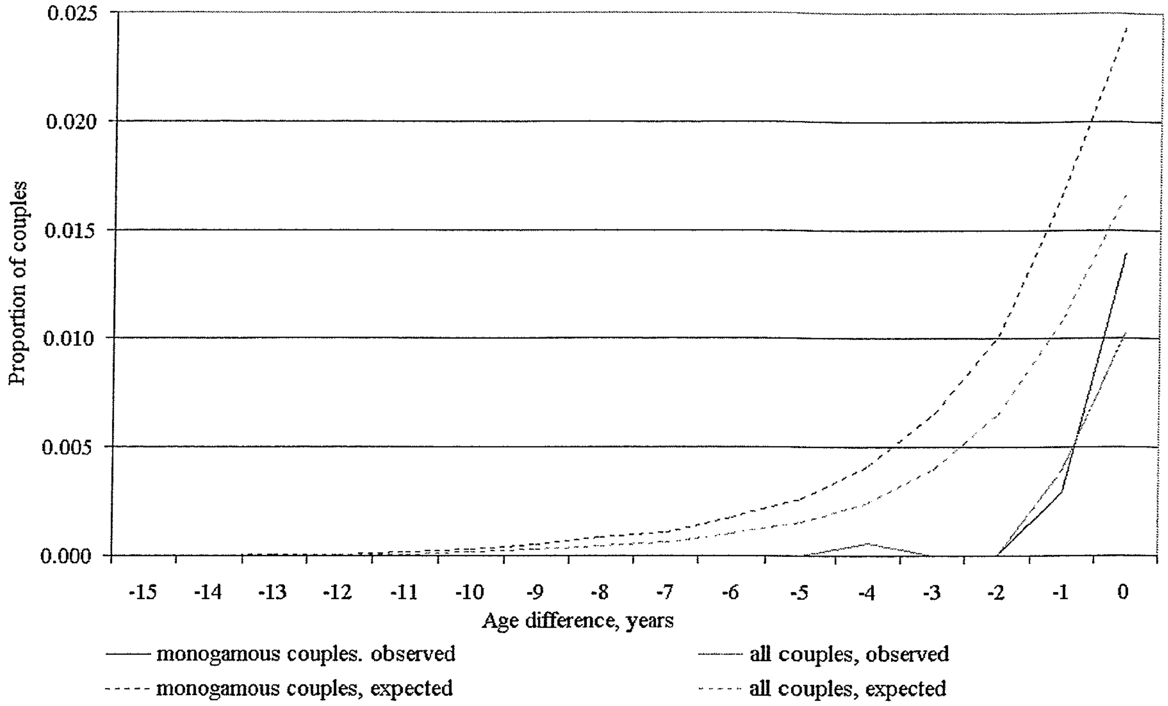
Figure 4.14 The observed and expected cumulative distributions of the age difference: The proportion of couples with an age difference of at least ten years in Haiti in comparison to Mali



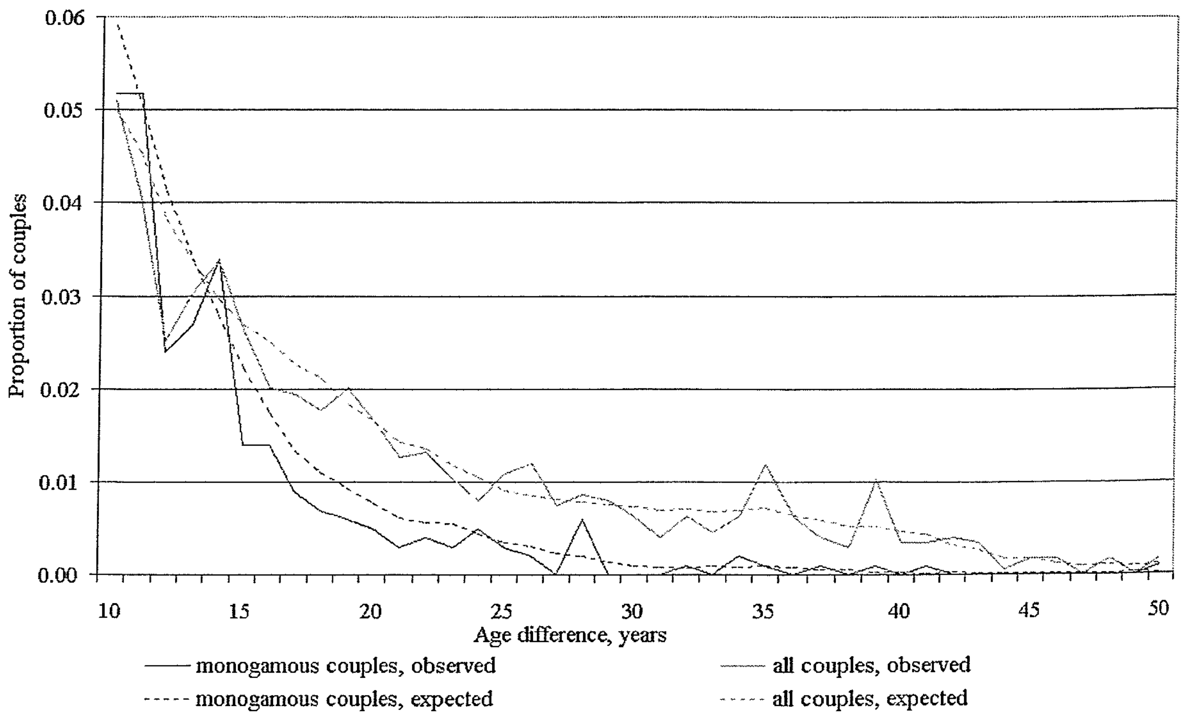
Generally, there is a larger difference between the observed and expected proportions of both 'extreme' age differences when focussing upon monogamous couples only, relative to the samples of all couples. Figure 4.15 illustrates this for Burkina Faso, the sample with smallest proportion of monogamous couples, by showing the observed and expected distributions of negative age differences and age differences of at least ten years, for the sample of all couples and the sample of just monogamous couples.

Figure 4.15 Selected sections of the observed and expected age difference distributions: Burkina Faso

Negative age differences



Age differences of 10+ years

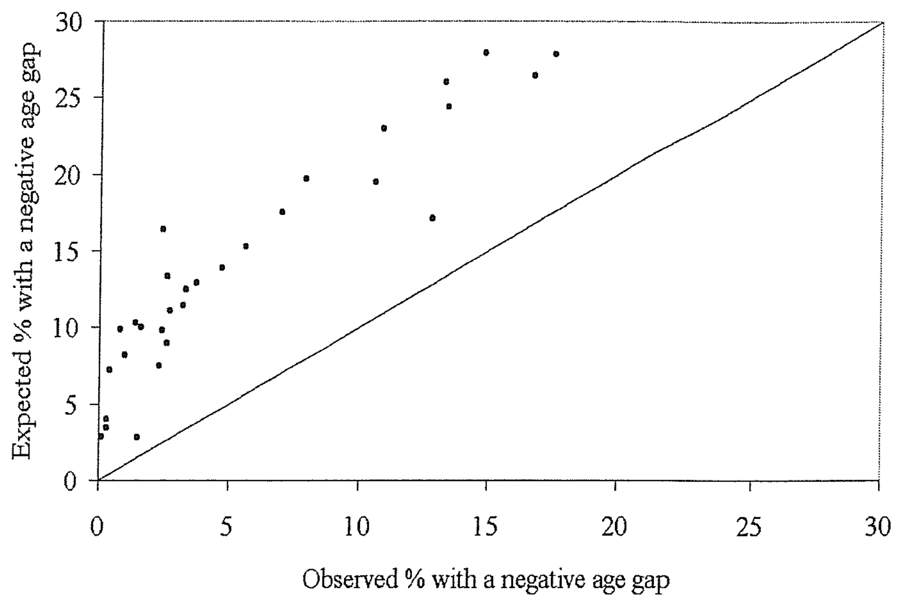


In contrast to these findings, Casterline *et al* (1986) reported relatively homogeneous estimates for the expected proportion of couples with more extreme age differences. For almost all of the countries studied by Casterline *et al* (1986), approximately one fifth of couples are expected to have a negative age difference. In comparison, the expected estimates calculated from the DHS data range from 2.8% of couples in Burkina Faso to 33.1% of couples in the Philippines, a range of over 30%. Similarly, Casterline *et al* (1986) estimated that the expected proportion of couples with an age difference of at least ten years ranges from 15% for Korea to 38% for Mauritania. This range is smaller than calculated from the DHS data for which, the expected proportion of couples with an age difference of at least ten years ranges from 5.8% for Uzbekistan to 56.7% for Mali, a difference of over 50%. Figures 4.16 and 4.17 show the differences between these estimates.

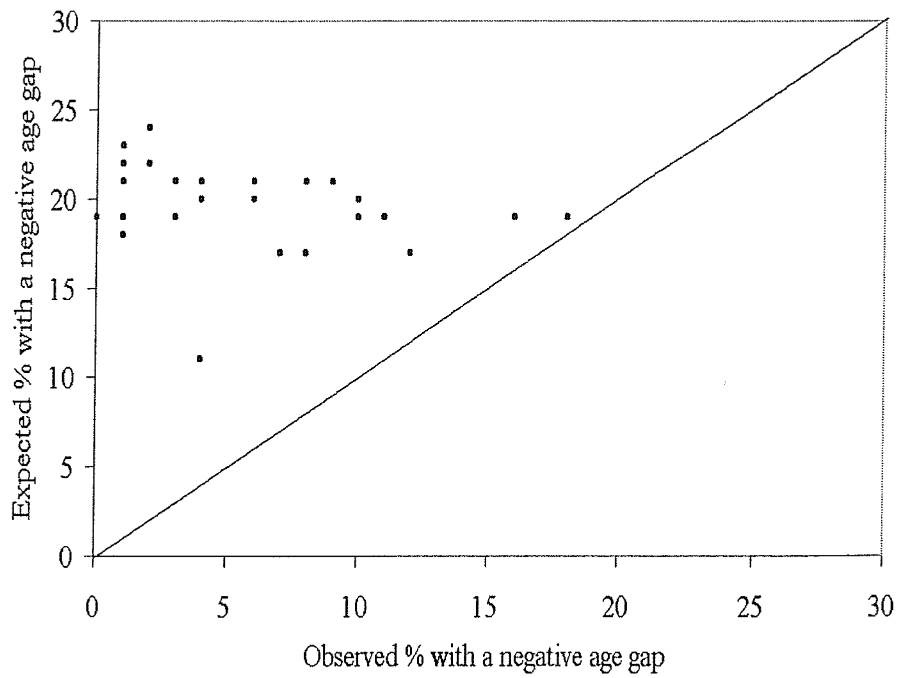
It is reasonable to expect the proportions with relatively more extreme age differences to be closely related to the observed age differences, thus, given this further disparity between the two studies, estimates of the observed and expected proportions with these more extreme age differences were calculated for the five WFS surveys mentioned above. Table 4.11 gives these estimates as well as those reported by Casterline *et al* (1986) for comparison. This re-analysis estimates that the expected percentages of couples with a negative age difference ranges from 2.1% for Bangladesh to 28.5% for the Philippines for these five WFS samples. There is thus, as much as 16.9 percentage points difference, for Bangladesh, between the expected percentages that Casterline *et al* (1986) report and those obtained from re-running their analyses. Similarly, while Casterline *et al* (1986) reported that between 19% and 33% of couples in these five countries are expected to have an age difference of at least ten years, in Egypt and Ghana respectively, the re-analysis estimates that between 16.5% and 48.5% of couples are expected to have such an age difference, in the Philippines and Bangladesh respectively. Thus, for both 'extreme' measures, the estimates reported by Casterline *et al* (1986) are more homogeneous than those obtained from re-analysing the WFS data. This is evident from Figure 4.18, which plots the observed and expected 'extreme' estimates for these five countries as reported by Casterline *et al* (1986) and as estimated from the re-analysis.

Figure 4.16 Scatterplots showing the association between the observed and expected proportions of all couples with a negative age difference: Comparing estimates from 30 DHS surveys with estimates from 28 WFS surveys as reported by Casterline *et al* (1986, p. 363)

30 DHS surveys



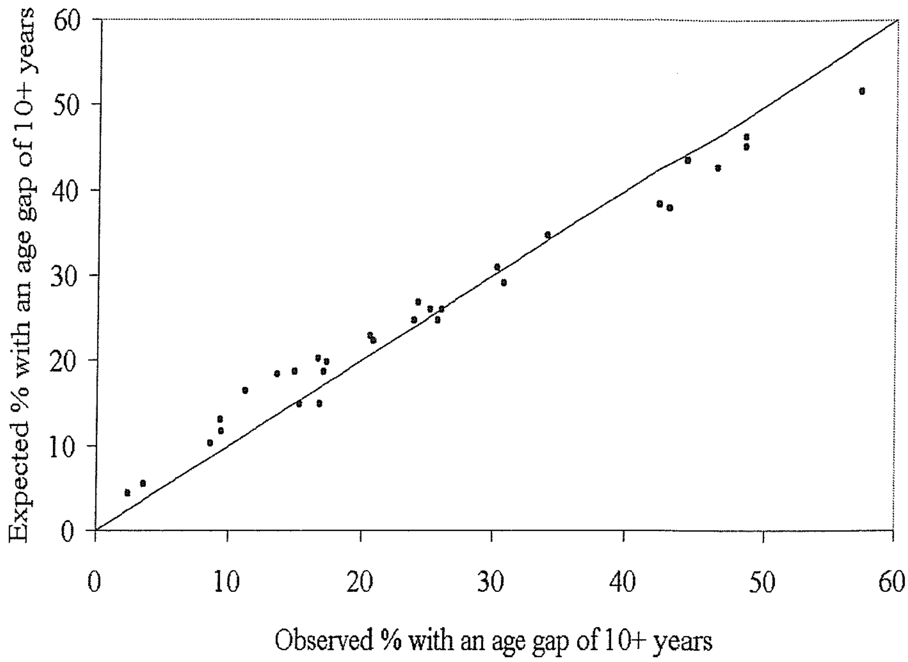
28 WFS surveys



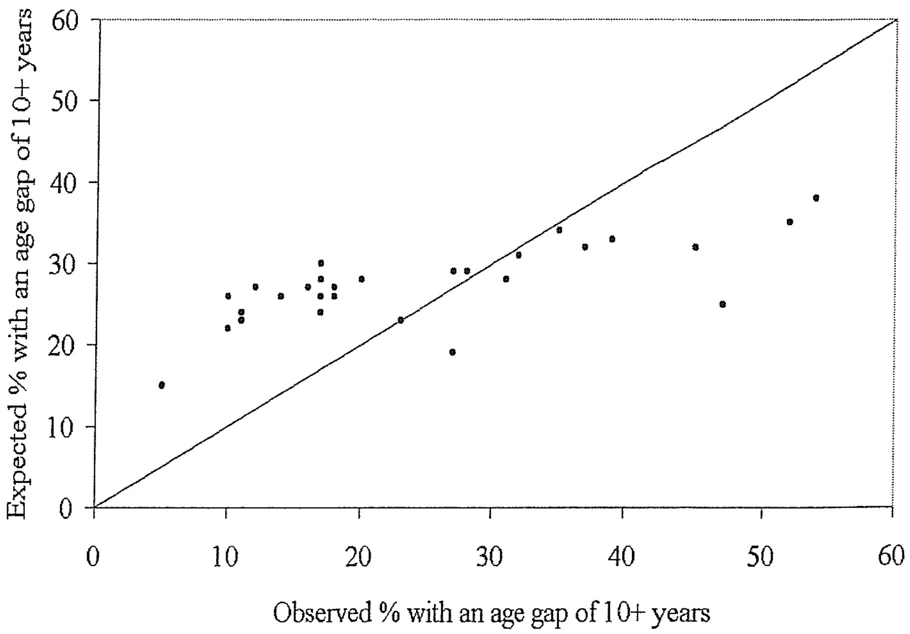
Note for Figure 4.16
WFS estimates as reported by Casterline *et al* (1986, p.363)

Figure 4.17 Scatterplots showing the association between the observed and expected proportions of all couples with an age difference of at least ten years: Comparing estimates from 30 DHS surveys with estimates from 28 WFS surveys as reported by Casterline *et al* (1986, p. 363)

30 DHS surveys



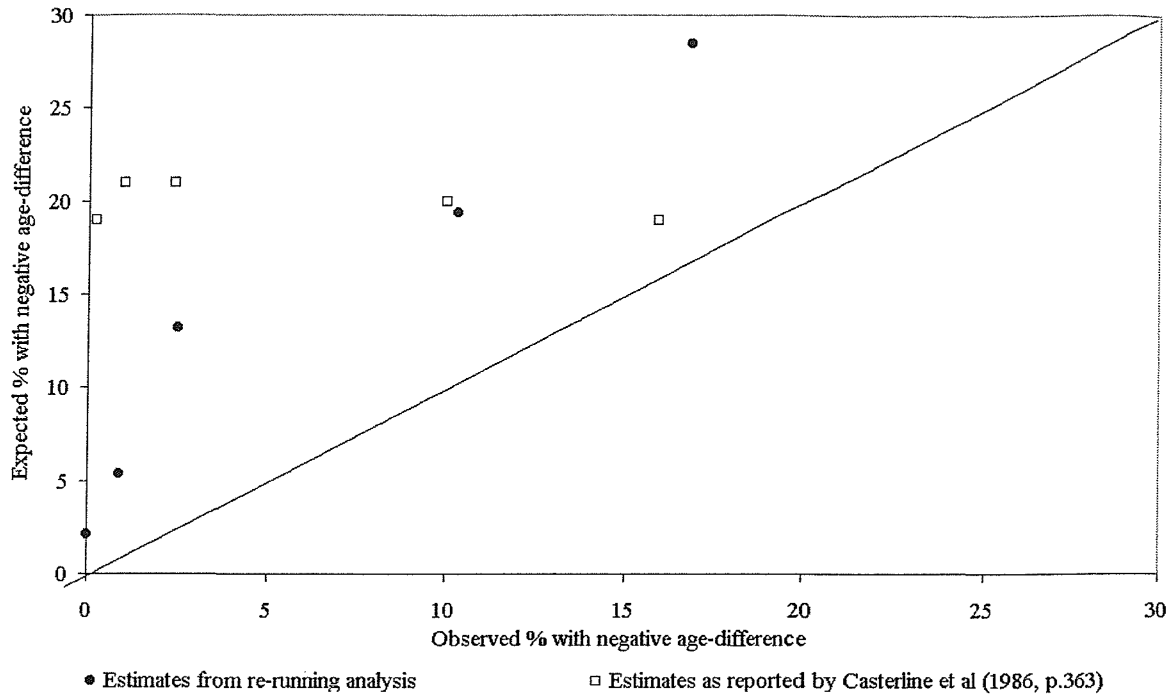
28 WFS surveys



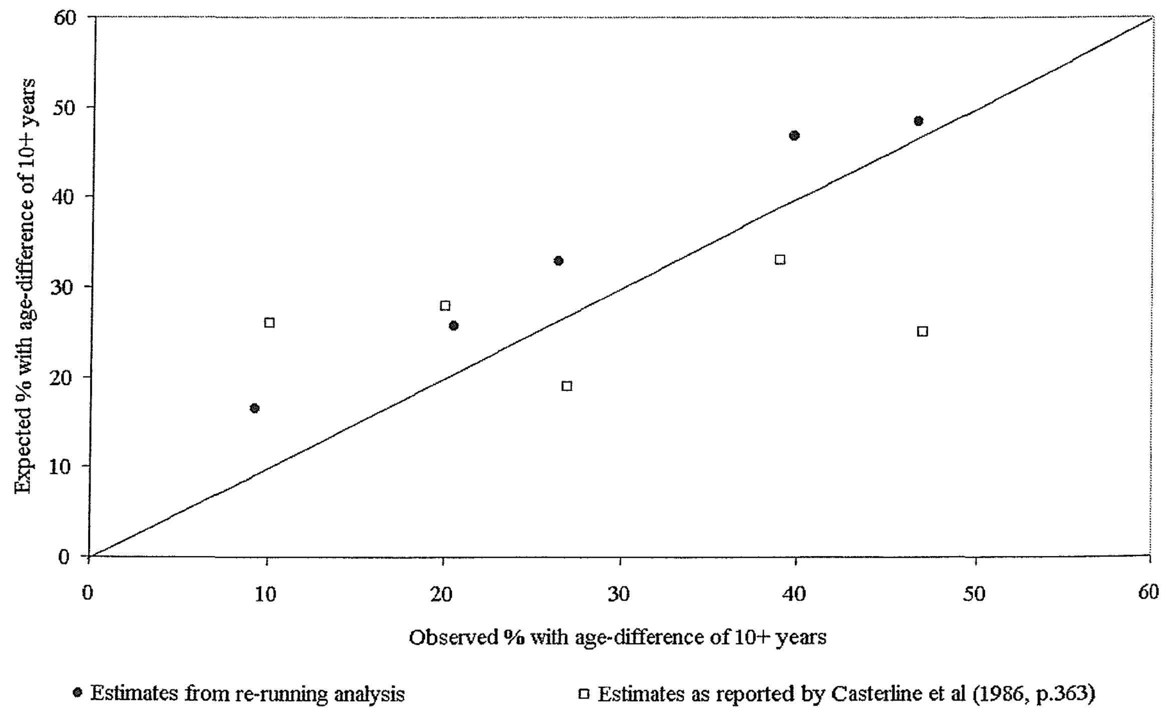
Note for Figure 4.17
WFS estimates as reported by Casterline *et al* (1986, p.363)

Figure 4.18 Scatterplots showing the association between the observed and expected percentages from the age difference distributions: Comparing estimates for five of the 28 WFS surveys reported by Casterline *et al* (1986) and from re-running their analysis

Percent of couples with a negative age difference



Percent of couples with an age difference of at least ten years



4.10 Discussion and conclusions

There were two main objectives of the analysis presented in this chapter. Firstly, to examine variations in the average age difference between countries and variations in the distributions of age differences within countries using individual-level data from as many DHS surveys as were available at the time of data collation. In total, 30 less developed countries were represented in this analysis using data that were collected between 1990 and 1997 as part of the DHS programme. This compares to a sample of 28 less developed countries that Casterline *et al* (1986) focussed upon in their analysis of data collected in the 1970s for the WFS programme. Unfortunately, only eleven countries appear in both samples and so comparisons are limited. Furthermore, this study acknowledges that the regional estimates may not be representative due to incomplete regional samples, as defined by the United Nations Statistical Division.

Despite this limited picture, extensive variation was observed in the average age difference both between and within the 30 countries, as observed in the aggregate level analysis in Chapter 3 and elsewhere (e.g. Casterline *et al*, 1986; United Nations, 1990; Ní Bhrolcháin, 1997). Age differences were found to be significantly larger, on average, in the sub-Saharan African countries studied, especially those in West Africa. Since up to 42.7% of women in the West African samples had a polygamous husband, and since age differences were found to be statistically significantly larger for such unions, it could be hypothesised that the larger average age differences observed in West Africa and in some sub-Saharan African countries reflects the prevalence of polygamy in these countries. Unlike the study by Casterline *et al* (1986), this chapter examined this hypothesis by additionally focussing solely on women married to monogamous husbands. However, average age differences in West Africa remained statistically significantly larger than elsewhere. One hypothesis maybe that, given the relatively high prevalence of polygamy in these societies and the tendency for larger age differences, it maybe considered desirable for monogamous couples to have larger age differences as well, with the associated implications for the timing of marriage for men and women. More generally, these results suggest that there is something unique about marriage patterns and the relative timing of marriage for men and women in West Africa relative to other regions of the world. Some attribute this to the continuing importance of kinship structure combined with the increasing role of individual partner selection in these societies (e.g. Lesthaeghe *et al*, 1989; Isiugo-Abanihe, 1994; Barbieri and Hertrich, 1999). As acknowledged by many authors in the demographic literature, union formation is often a process rather than a single event and so it is difficult to understand these processes through observational data such as the DHS survey data (e.g. van de Walle, 1968; United Nations, 1990; Meekers, 1992; Isiugo-Abanihe, 1994). Furthermore, as noted in Chapter 3, it is recognised that these processes vary by culture and so analyses conducted at the aggregate level are limited in their generalisability (Obermeyer, 1992).

Throughout this chapter, care has been taken to discuss cross-national variations in the age difference in terms of the average age difference, since much variation in the distribution of the age difference was observed for all countries, especially those with a larger average age difference. In addition to considering cross-national and within-country variations in the average age difference, this analysis also looked at cross-national variations in the average age at marriage and the distributions of age at marriage within countries for the men and women in these samples. As observed in Chapter 3, greater variation was observed between countries in the average age at marriage of females rather than males, while greater variation was observed within countries in the age at marriage of husbands rather than wives, as Casterline *et al* (1986) also observed. This result was observed for the samples of all couples and also for the samples of just monogamous couples, so this result can not be attributed to the greater age range of polygamous husbands relative to monogamous husbands. Also, while the sample selection criteria used for this analysis only applied to women, these did not unduly constrain the data with regard to the age range that the majority of women in these and in fact, most societies marry within (e.g. United Nations, 1990).

Secondly, this chapter sought to use the DHS data to examine one of the hypotheses proposed in the literature to explain these variations, that union formation occurs randomly with respect to age and simply reflects the age-sex structure of those eligible for marriage, i.e. in the marriage market (e.g. Casterline *et al*, 1986; Ní Bhrolcháin, 2000). Of course, it is important to acknowledge that the marriage market hypothesis could not be thoroughly investigated since the DHS data refer to individuals who were in union, rather than all eligible men and women in the marriage market. In addition, the extent to which these samples represent the eligible population could not be ascertained, not least because the marriage market is not limited by geographical boundaries (Smith, 1983). Despite these limitations, the hypothesis was examined using joint distribution theory to compare the observed distribution of age differences with the expected distribution of age differences, that is, the distribution that would be achieved if the couples in the sample had formed randomly with respect to age. This is the approach that Casterline *et al* (1986) used but this chapter's analysis of the DHS data discovered very different results. The medians observed from the DHS data tended to be approximately as expected, which were considered as evidence in support of the hypothesis that mating occurs randomly with respect to age, at least on average. However, Casterline *et al*'s (1986) expected medians were around five to six years for all countries studied, and so they consequently rejected this hypothesis. It is assumed that the authors would have reviewed their methods and results, especially since it seems logical to expect a relationship between the average age difference and the average age at marriage of men and women, and thus it seems reasonable to assume that the expected median age differences would be strongly related to the observed median age differences. However, re-analysis of data from five of the WFS surveys that Casterline *et al* (1986) studied produced results consistent with those observed from the analysis of the DHS data. It is therefore necessary to question their calculations and thus their results.

While evidence was found in support of the random mating hypothesis as far as the average age difference was concerned, all of the DHS samples showed evidence of a concentration of the age difference distributions around the average and an increase in the difference between the observed and expected distributions with increasing distance from the average. Unlike the results referring to the average age difference, these results are consistent with Casterline *et al*'s (1986) analysis. Furthermore, for all sample distributions the difference between the observed and expected proportions was larger for the proportion of couples with a negative age difference than the proportion of couples with an age difference of at least ten years. Thus, even when age differences were observed as small on average, negative age differences - while not uncommon - were observed much less frequently than expected. Thus, while the hypothesis that mating occurs at random with respect to age applies on average, this hypothesis clearly does not apply to the distributions in their entirety. Some other explanation is therefore necessary. Casterline *et al* (1986) interpreted this as evidence of a tendency to avoid certain age differences in all of the countries that they studied, and proposed that the results were *"strong evidence of the operation of age difference preferences"* (Casterline *et al*, 1986, p. 366). While this explanation *may* apply, this chapter is hesitant to draw such a

conclusion since, as discussed in Chapter 1, the preferences hypothesis can only be examined with explicit preferences data (Ní Bhrolcháin, 2000).
This chapter concludes therefore that there is some evidence that the observed distributions of age differences reflect the age-sex structure of those marrying, at least on average, such that union formation occurs randomly with respect to age, on average. However, the smaller observed than expected IQRs for the age difference seem to reflect a tendency towards matching by age. This could be based on ordinal matching rather than being a tendency towards exact matching, and so could be independent of the difference between the average ages at marriage of the two sexes. Furthermore, the relative skewness of the observed distributions of age differences in comparison to the expected distributions suggests the existence of an asymmetry in the underlying process, and thus something over and above structure and ordinal matching, such as the existence of a stigma attached to negative differences.

The age-sex structure of those marrying is important, and the case for the existence of age preferences is unproven, but the empirical findings clearly reflect more than straightforward consequences of the age-sex structure. As discussed in Chapter 1, the hypothesis that variations in the age difference are a by-product of factors associated with age at marriage may also apply as this is a complementary hypothesis. Thus, the next chapter examines this 'by-product hypothesis'.

Chapter 5:
Results of examining the 'by-product' hypothesis that within-country variations in the age difference are by-products of factors associated with age at marriage

5.1 Overview

One of the observations made in Chapter 4 was that within countries, there is relatively more variation in the distribution of the husband's age at marriage than in the distribution of the wife's age at marriage. Furthermore, the amount of variation in the husband's age distribution relative to the wife's age distribution varies between countries. In general, the greater variation in the husband's age at marriage relative to the wife's age, the greater the proportion of the variation in the age difference that can be accounted for by the husband's age at marriage. However, as Casterline *et al* (1986) observed, the greater the variation in the husband's age relative to the wife's age, the smaller the magnitude of association between the spouses' ages. Nevertheless, an association between the spouses' ages exists in all countries.

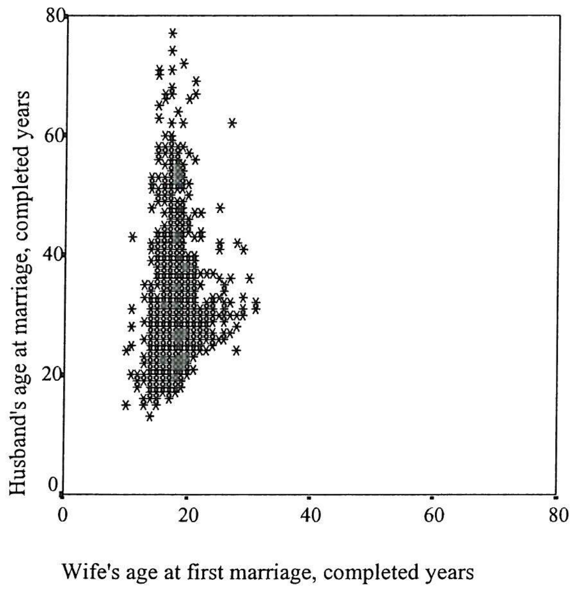
This chapter uses individual-level data collected by the DHS surveys for each of the 30 countries studied in Chapter 4 to investigate the hypothesis that the association between the ages at marriage of husbands and wives is a by-product of associations between factors associated with the ages at marriage of husbands and wives. Thus, this hypothesis assumes that the pattern of association between the husband's age and the wife's age is responsible for the systematic component of the distribution of age differences, i.e. the difference between the observed distribution and the distribution expected given the 'random' matching of spouses with respect to their ages at marriage. To begin, the extent of association between the ages at marriage of husbands and wives is examined for each country. The by-product hypothesis is then examined by calculating the partial correlation coefficient for the association between the ages at marriage of husbands and wives net of a number of factors associated with age at marriage, which is an extension of the approach adopted by Casterline *et al* (1986). However, this method only considers marginal associations so the chapter proceeds to use hierarchical interaction models to examine the association between the ages at marriage of husbands and wives net of both marginal and conditional associations with factors associated with age at marriage (section 2.3.7.5 illustrates the difference between marginal and conditional associations).

5.2 The association between the ages at marriage of husbands and wives

There is greater within-country variation in the distribution of the husband's age at marriage relative to the distribution of the wife's age at marriage, yet it is important to recall from Chapter 4 that variations in the wife's age at marriage do exist albeit to a lesser degree. Furthermore, as Table 5.1 shows there is an association between the age at marriage of wives and the age at marriage of husbands within countries. Table 5.1 gives the Pearson correlation coefficients as a measure of this association for each of the 30 samples for all couples, and also just for monogamous couples. In all cases, the Pearson correlation coefficient is positive signifying a positive association between the ages at marriage of husbands and wives. The Pearson correlation coefficient ranges from .08 for the sample of all couples in Burkina Faso to .67 for the sample of all couples in Uzbekistan. As examples, Figure 5.1 shows scatterplots of the husband's age at marriage plotted against the wife's age at marriage for these two countries. The difference between the two countries in terms of the amount of association is evident from Figure 5.1, as is the difference between the two countries in terms of the amount of variation in the distributions of ages at marriage of husbands and wives.

Figure 5.1 Scatterplots of the husband's age at marriage plotted against the wife's age at marriage for the two countries with the smallest and largest Pearson correlation coefficients of the sample of 30 countries

Burkina Faso (all couples)
Pearson correlation coefficient: .08



Uzbekistan (all couples)
Pearson correlation coefficient: .67

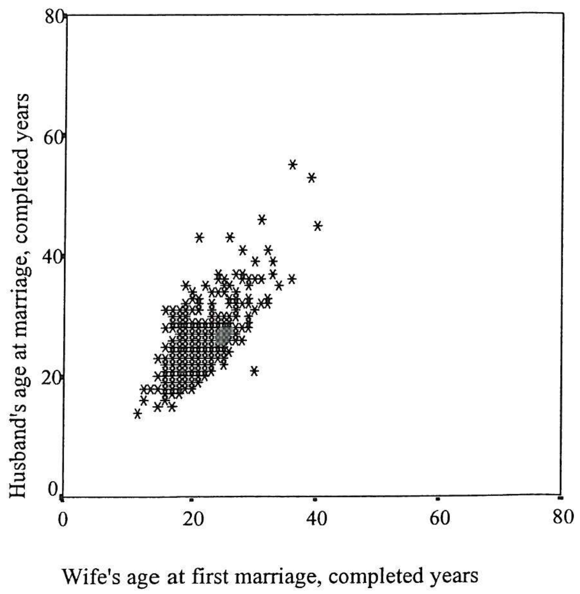


Table 5.1 The Pearson correlation coefficient [1] as a measure of the association between the ages at marriage of husbands and wives, by country [2]

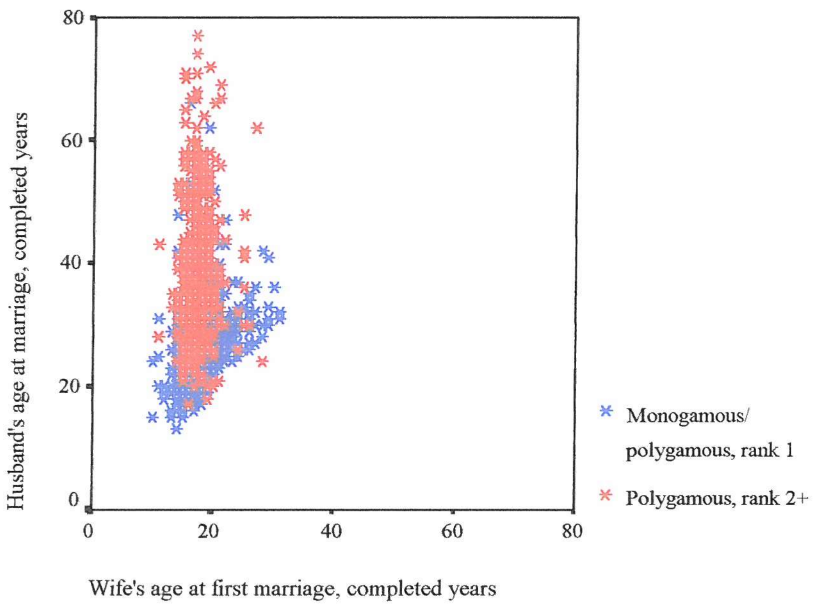
Region and country	All couples	Monogamous couples only
Central & South America		
Guatemala	.45	.45
Peru	.56	.56
Brazil	.48	.48
Colombia	.50	.50
Dominican Republic	.40	.40
Haiti	.34	.34
West Asia & North Africa		
Kazakhstan	.61	.61
Uzbekistan	.67	.67
Egypt	.49	.49
South Asia		
Philippines	.48	.48
Nepal	.50	.52
Pakistan	.49	.52
Bangladesh	.38	.38
Central, East & South Africa		
Rwanda	.23	.26
Uganda	.43	.41
Malawi	.42	.47
Mozambique	.46	.52
Central African Republic	.39	.41
Zambia	.55	.55
Zimbabwe	.34	.47
Kenya	.35	.42
Tanzania	.37	.47
Comoros	.41	.41
West Africa		
Ghana	.43	.49
Benin	.26	.37
Niger	.26	.30
Burkina Faso	.08	.29
Cameroon	.24	.34
Côte d'Ivoire	.26	.42
Mali	.27	.37

Notes for Table 5.1

1. All Pearson correlation coefficients are statistically significant at $p<.01$
2. Countries are presented in order of magnitude of the observed median age difference (Table 4.1)

Table 5.1 shows that, excluding Burkina Faso, the next smallest correlation between all husbands and wives is observed for Rwanda at .23. In general, there is a negative association between the size of the correlation coefficient and the median age difference. Thus, the correlation coefficient tends to be smaller for the West African samples where, on average, age differences tend to be relatively large (as identified in Chapter 4), in comparison to samples where age differences tend to be relatively small, on average, as in Uzbekistan. This follows since, as Chapter 4 showed, samples with larger average age differences tend to have a greater variation in the distribution of the husband's age at marriage relative to the wife's age at marriage, and as such, a smaller association between the spouses' ages. It follows therefore that for the surveys that collected data on polygamy the correlation coefficient tends to be larger for the samples of monogamous couples relative to the samples of all couples since there tends to be less variation in the husband's age at marriage, and thus a greater association between the spouses' ages. Figure 5.2 illustrates this by plotting the husband's age at marriage against the wife's age at marriage and identifying monogamous couples or polygamous couples where the wife is rank 1, and polygamous couples in which the wife is ranked at least 2 for Burkina Faso, the sample with the greatest proportion of polygamous couples where the wife is not the first wife (see Appendix 4A). The corresponding Pearson correlation coefficients for these two sub-samples are respectively, .23, which is highly statistically significant (p=.000) and .005, which is not statistically significant (p>.05).

Figure 5.2 Scatterplots of the husband's age at marriage plotted against the wife's age at marriage for Burkina Faso: Comparing the association for monogamous couples and couples where the wife is rank 1 with polygamous couples where the wife is rank 2+



Note for Figure 5.2
Some of the red asterisks denoting 'polygamous, rank 2+' overlay the blue asterisks denoting 'monogamous/polygamous, rank 1'

The hypothesis under investigation in this chapter is whether the observed association between the husband's age at marriage and the wife's age at marriage is a by-product of factors associated with the timing of marriage at the individual-level. The by-product hypothesis is initially examined by calculating the partial correlation coefficient. This is a measure of the extent of linear association between the husband's age at marriage and the wife's age at marriage while controlling for the linear effects of one or more additional variables. The next section considers the variation in the median age at marriage for husbands, and then wives, for each of the factors available from the 1995 Egyptian DHS survey, as an example.

5.3 Factors hypothesised as associated with age at marriage

Chapter 2 identified a number of factors that are hypothesised to be associated with age at marriage that are available from the DHS surveys (Table 2.4). Table 5.2 gives the median age at marriage for husbands and wives (and the 95% confidence interval) for seven factors for which data are available for the Egyptian sample of couples. There are highly statistically significant differences in the median age at marriage of both husbands and wives for all but one of these factors (p=.000). For this sample, the statistically significant differences are in the same direction for both husbands and wives.

Table 5.2 shows that there is a positive association between age at marriage and household socio-economic status, wife's education, and husband's education, for both husbands and wives. Thus, for example, a husband without education married 4.9 years earlier on average than a husband with more than secondary education at 26.0 years and 30.9 years respectively. Similarly, a wife without education married 6.3 years earlier on average than a wife with more than secondary education at 18.7 years and 25.0 years respectively. In contrast, there is a negative association between age at marriage and consanguinity. On average, marriage occurred approximately two to three years earlier among respondents in rural areas or when the wife works for cash. These results support the evidence discussed in Chapter 1. The one exception for this sample of Egyptian couples is religion for which there is no statistically significant difference in the median age at marriage of husbands (p>.05 according to a t-test), although marriage for Muslim wives occurred one year earlier on average than for Egyptian women who are not Muslim, which is highly statistically significantly (p=.000 according to a t-test).

Table 5.2 Variations in the median ages at marriage of husbands and wives for factors hypothesised to be associated with age at marriage for a sample of couples from the 1995 Egyptian DHS

Factor and categories	% of weighted sample (n=5343)	Age at marriage of: Husbands Median (95% CI)	Age at marriage of: Wives Median (95% CI)
Place of residence		*p*=.000 [1]	*p*=.000 [2]
rural	53.5	26.1 (25.9-26.3)	19.2 (19.1-19.3)
urban	46.5	28.7 (28.5-29.0)	22.0 (21.8-22.2)
	(100.0)		
Household socio-economic status		*p*=.000 [4]	*p*=.000 [4]
low	32.4	26.2 (25.9-26.5)	19.1 (18.9-19.3)
average	45.0	26.9 (26.7-27.2)	20.1 (20.0-20.3)
high	22.6	29.8 (28.5-30.1)	23.3 (23.1-23.6)
	(100.0)		
Wife earns cash for work		*p*=.000 [2]	*p*=.000 [1]
no	85.5	27.0 (26.8-27.2)	19.9 (19.8-20.0)
yes	14.5	29.5 (29.1-29.8)	24.2 (23.9-24.5)
	(100.0)		
Wife's education		*p*=.000 [4]	*p*=.000 [4]
none	31.1	26.0 (25.6-26.3)	18.7 (18.6-18.9)
primary	16.3	26.6 (26.2-27.0)	19.8 (19.5-20.1)
secondary	43.2	27.9 (27.7-28.1)	21.1 (20.9-21.2)
higher	9.3	30.9 (30.5-31.2)	25.0 (24.7-25.3)
	(100.0)		
Husband's education		*p*=.000 [4]	*p*=.000 [4]
none	16.9	26.7 (26.2-27.2)	19.0 (18.7-19.3)
primary	23.5	26.6 (26.2-26.9)	19.5 (19.3-19.8)
secondary	43.7	27.0 (26.8-27.2)	20.5 (20.4-20.7)
higher	15.8	30.1 (29.8-30.4)	23.5 (23.3-23.8)
	(100.0)		
Religion		n.s. [1]	*p*=.000 [1]
Muslim	94.5	27.3 (27.2-27.5)	20.4 (20.3-20.6)
else (predominantly Christian)	5.5	27.7 (27.1-28.2)	21.5 (21.0-22.0)
	(100.0)		
Consanguineous marriage		*p*=.000 [4]	*p*=.000 [4]
first cousin	21.3	25.5 (25.3-25.8)	19.4 (19.2-19.6)
second cousin/other relative	16.5	26.6 (26.3-27.0)	19.4 (19.2-19.7)
not related	62.3	28.2 (27.9-28.4)	21.2 (21.0-21.3)
	(100.0)		

Notes for Table 5.2

Statistical significant variation in the median age at marriage tested with:

1. t-test as variance homogeneity can be assumed
2. Mann-Whitney test as variance homogeneity can not be assumed
3. ANOVA as variance homogeneity can be assumed
4. Kruskal-Wallis test as variance homogeneity can not be assumed

5.4 The partial correlation between the ages at marriage of husbands and wives

Given the hypothesised and observed associations between the factors identified in Table 2.4 and the ages at marriage of husbands and wives, this section examines the partial correlation coefficient between the ages at marriage of husbands and wives net of the available factors, for each of the 30 countries, for all couples and also just for monogamous couples. Table 5.3 shows these partial correlation coefficients in bold, together with the Pearson correlation coefficients, which are given for reference.

Table 5.3 shows that for all samples there is little, if any difference between the partial correlation coefficient and the Pearson correlation coefficient. This can be interpreted as indicating that there is little or no change in the magnitude of the association between the two ages at marriage net of the linear effects of the available factors that are hypothesised as associated with age at marriage. The greatest change is observed for the samples of all couples in Burkina Faso and Côte d'Ivoire, where the partial correlation coefficient is .07 points larger than the Pearson correlation coefficient.

The partial correlation coefficient tends to be larger than the Pearson correlation coefficient for countries where age differences tend to be larger, on average. However, when this analysis is repeated for the samples of monogamous couples, the partial correlation coefficient tends to be smaller than the Pearson correlation coefficient. In general, there does not seem to be any regional patterns in the magnitude of the change in the correlation coefficient. Of course, it is important to recall that the partial correlation coefficients for the 30 samples are calculated net of the linear effects of different factors, reflecting the data available from the corresponding DHS survey (Table 2.4).

Table 5.3 Pearson and partial correlation coefficients as measures of the association between the ages at marriage of husbands and wives for all couples and for monogamous couples only, by country

Sample:	All couples		Monogamous couples only	
Correlation coefficient:	Pearson [1]	Partial [1,2]	Pearson [1]	Partial [1,2]
Region and country [3]:				
Central & South America				
Guatemala	.45	**.42**	.45	**.42**
Peru	.56	**.53**	.56	**.53**
Brazil	.48	**.48**	.48	**.48**
Colombia	.50	**.49**	.50	**.49**
Dominican Republic	.40	**.37**	.40	**.37**
Haiti	.34	**.39**	.34	**.39**
West Asia & North Africa				
Kazakhstan	.61	**.61**	.61	**.61**
Uzbekistan	.67	**.66**	.67	**.66**
Egypt	.49	**.43**	.49	**.43**
South Asia				
Philippines	.48	**.47**	.48	**.47**
Nepal	.50	**.47**	.52	**.48**
Pakistan	.49	**.49**	.52	**.50**
Bangladesh	.38	**.36**	.38	**.36**
Central, East & South Africa				
Rwanda	.23	**.21**	.26	**.22**
Uganda	.43	**.38**	.41	**.36**
Malawi	.42	**.41**	.47	**.46**
Mozambique	.46	**.45**	.52	**.52**
Central African Republic	.39	**.40**	.41	**.41**
Zambia	.55	**.51**	.55	**.51**
Zimbabwe	.34	**.39**	.47	**.48**
Kenya	.35	**.39**	.42	**.43**
Tanzania	.37	**.37**	.47	**.47**
Comoros	.41	**.35**	.41	**.40**
West Africa				
Ghana	.43	**.45**	.49	**.48**
Benin	.26	**.27**	.37	**.30**
Niger	.26	**.21**	.30	**.23**
Burkina Faso	.08	**.15**	.29	**.26**
Cameroon	.24	**.29**	.34	**.32**
Côte d'Ivoire	.26	**.33**	.42	**.42**
Mali	.27	**.27**	.37	**.33**

Notes for Table 5.3

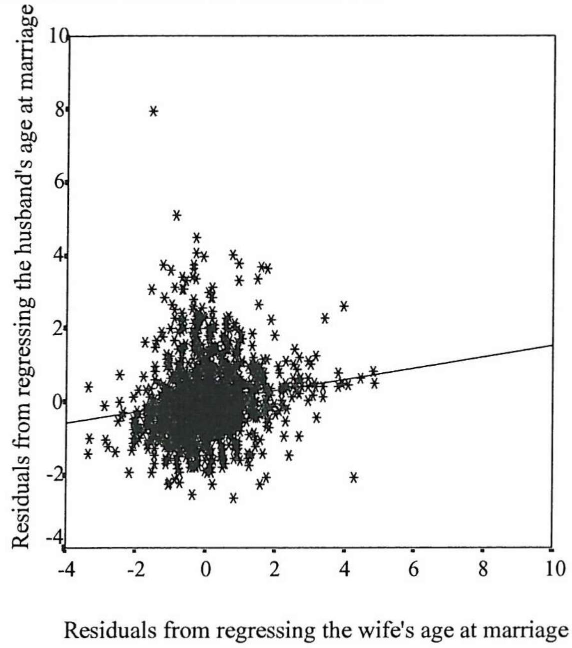
1. All Pearson and partial correlation coefficients are statistically significant at $p<.01$
2. Net of the available factors associated with age at marriage available identified in Table 2.4
3. Countries are presented in order of magnitude of the observed median age difference (Table 4.1)

The partial correlation coefficients given in Table 5.3 are now displayed graphically for Burkina Faso and Uzbekistan, the countries with respectively, the largest and smallest partial correlation coefficients (and the largest and smallest Pearson correlation coefficients, as identified in Table 5.1).

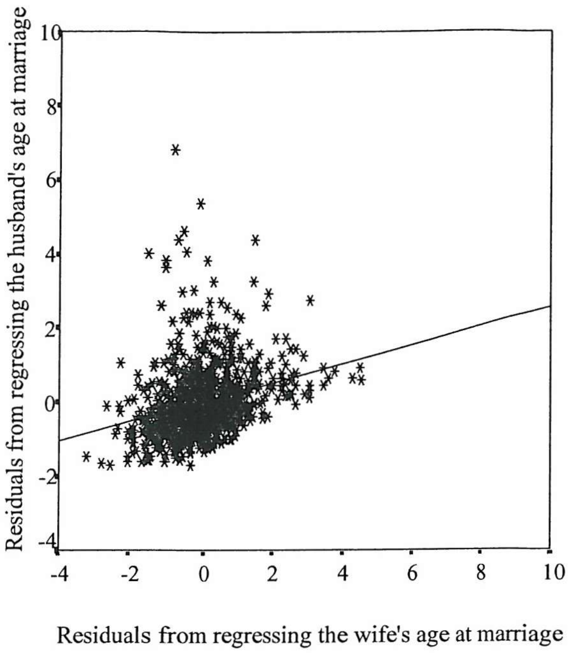
Figure 5.3 shows for Burkina Faso, the association between the husband's age at marriage and the wife's age at marriage net of the linear effects of the factors associated with age at marriage identified in Table 2.4, firstly for all couples and then just for monogamous couples. The scatterplots are achieved by plotting the residuals obtained from regressing the husband's age at marriage on these factors against the residuals obtained from regressing the wife's age at marriage on the same factors. The least-squares linear regression line that best fits the data is also displayed. These regression lines identify a positive association in all three plots in Figure 5.3. Thus, where the factors that are hypothesised as associated with age at marriage over-estimate the wife's age at marriage, it is likely that the same factors over-estimate the husband's age at marriage. However, these associations are far from perfect with much variation in the size of the residuals in both directions for all plots.

Figure 5.3 Scatterplots of the residuals from regressing the husband's age at marriage on factors associated with age at marriage[1] and the residuals from regressing the wife's age at marriage on the same factors associated with age at marriage[1]

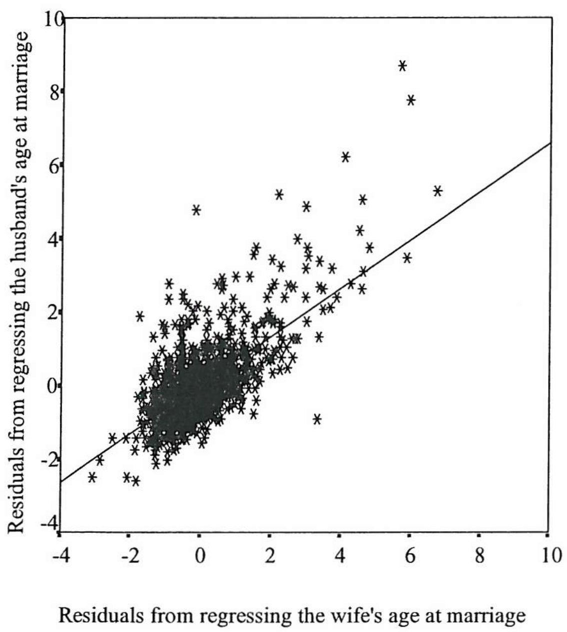
Burkina Faso
All couples, n=1642
Partial correlation coefficient: .15



Monogamous couples only, n=1027
Partial correlation coefficient: .26 [2]



Uzbekistan
All couples, n=1361
Partial correlation coefficient: .66



Notes for Figure 5.3

1. See Table 2.4 for the factors associated with age at marriage available for these samples
2. Neither age at marriage is regressed on the polygamy factor to obtain the residuals for the plot for the sample of just monogamous couples as this is constant for this sample.
3. Regression diagnostics were examined to ensure the assumptions of linear regression were met

A limitation of using the partial correlation coefficient and partial residuals is that these methods rely on pairwise marginal associations, which can result in spurious associations being identified (see section 2.3.7.5). Consequently, it is necessary to take a multivariate approach that includes all the relevant variables in the analysis in order to study the conditional as well as marginal associations. This is the basis of graphical modelling as discussed in Chapter 2. The next section presents the results of using this approach to examine whether there is a direct association between the ages at marriage of husbands and wives net of the conditional and marginal associations with the factors hypothesised as associated with age at marriage.

Given space limitations it is not possible to present the results for all 30 countries, especially since the analysis is also conducted separately for monogamous couples for surveys that collect data on polygamy. Thus, in addition to Egypt for which there is a particularly rich dataset (see section 2.3.7.4), the results for five contrasting countries are presented -one country from each of the five UN regions studied. However, it is acknowledged that these countries are not necessarily representative either in terms of their UN region or the number or types of factors for which data are available. The countries are: Uzbekistan, the sample with the greatest association between the ages at marriage of husbands and wives according to the Pearson and partial correlation coefficients; Burkina Faso, the sample with the smallest association between the ages at marriage of husbands and wives according to these correlation coefficients and a sample that collects data on polygamy; the Central African Republic, one of two samples that has data available for all ten factors and a sample that collects data on polygamy; Peru, one of the samples with the least number of factors available; and Bangladesh, a sample with a large average age difference that is not in sub-Saharan Africa.

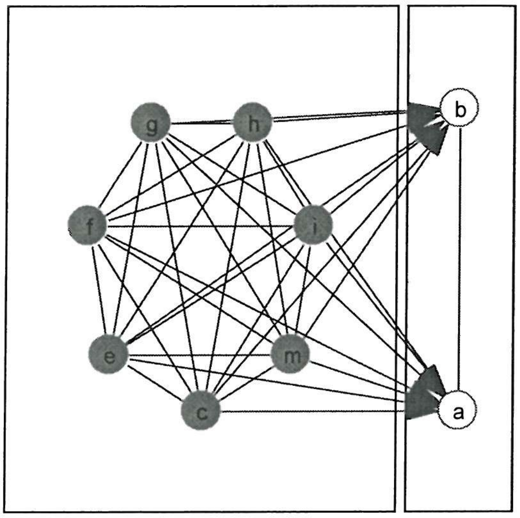
5.5 Graphical modelling the association between the ages at marriage of husbands and wives

The graphical modelling package MIM version 3.1 is used for the following analyses. MIM assigns a letter code to identify each variable (section 2.3.7.4 gives the categories for these variables). For all models, the variables are coded:

a	husband's age at marriage	h	husband's highest level of education
b	wife's age at marriage	i	religion
c	place of residence	j	ethnicity
d	married or cohabiting	l	household structure
e	household socio-economic status	m	consanguineous union
f	wife works and earns cash	o	polygamous husband
g	wife's highest level of education		

Thus, the saturated model for the Egyptian dataset where data are available for factors c, e, f, g, h, i and m, is expressed as cefghim | ab. This means that all the available variables are mutually dependent regardless of the blocking, that is, considering the associations among the factors associated with age at marriage *before* the associations between these factors and the ages at marriage of husbands and wives. Figure 5.4 shows this model graphically.

Figure 5.4 The saturated graphical chain model for the Egyptian sample



Key for Figure 5.4

a	husband's age at marriage
b	wife's age at marriage
c	place of residence
e	household socio-economic status
f	wife works and earns cash
g	wife's highest level of education
h	husband's highest level of education
i	religion
m	consanguineous marriage

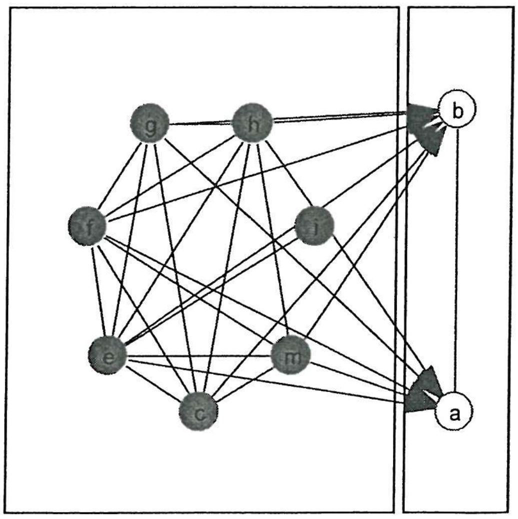
In a graphical model, the nodes represent the variables. Dots represent discrete variables (i.e. all the explanatory variables in this model) and circles represent continuous variables (i.e. the outcome variables in this model). The lines connecting variables are called 'edges' in MIM and represent a direct association between the two variables. Thus, where nodes are not connected, the variables they represent are considered conditionally independent given the other variables in the block and any preceding block.

Using the Egyptian data as an example, a parsimonious hierarchical interaction model is identified from the stepwise selection procedure ($p<.05$). In MIM notation, this is expressed as:

1. ei,cefhm,cefgh
2. cefghim/bcefghm,aefghm/ab

Thus, in block 1, there is a high degree of association among the factors hypothesised as associated with age at marriage. There is a direct association between household socio-economic status (e) and religion (i). Pairwise direct associations also exist between place of residence (c), household socio-economic status (e), whether the wife works and earns cash (f), the husband's education (h), and consanguinity (m). Also, between place of residence (c), household socio-economic status (e), whether the wife works and earns cash (f), the wife's education (g), and the husband's education (h). Net of the associations between the seven factors associated with age at marriage (block 1), all of the factors except religion and place of residence have a direct association with the husband's age at marriage and all except religion with the wife's age at marriage (block 2). These associations are shown as a graphical model in Figure 5.5.

Figure 5.5 A parsimonious graphical chain model of the associations between the ages at marriage of husbands and wives and factors associated with age at marriage: Egypt



Key for Figure 5.5

a	husband's age at marriage	g	wife's highest level of education
b	wife's age at marriage	h	husband's highest level of education
c	place of residence	i	religion
e	household socio-economic status	m	consanguineous marriage
f	wife works and earns cash		

It is not the purpose of this analysis to consider the direction of the association between these factors and age at marriage, rather it is to consider whether there remains a direct association between the two ages at marriage net of the other associations. As Figure 5.5 shows, in this parsimonious model there remains an edge, and thus a direct association between the ages at marriage of husbands and wives. This edge is identified as highly statistically significant (p=.000) from the edge exclusion test (see section 2.3.7.5).

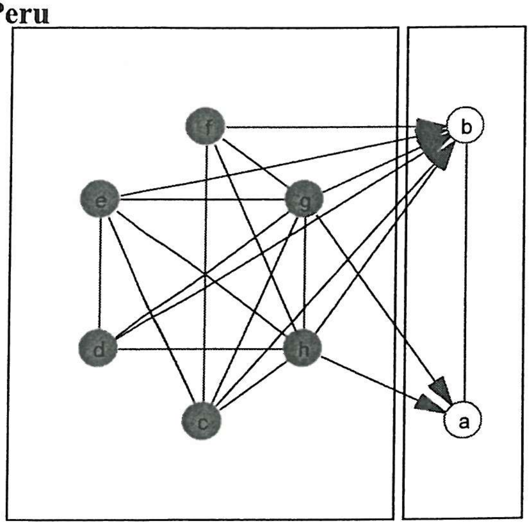
Figure 5.6 shows a parsimonious graphical chain model for each of the five contrasting countries identified above: Peru, Uzbekistan, Bangladesh, Central African Republic, and Burkina Faso. For each of the two sub-Saharan African samples, two parsimonious graphical models are shown: one for the sample of all couples and one for the sample of only monogamous couples.

Figure 5.6 Parsimonious graphical chain models for five contrasting countries

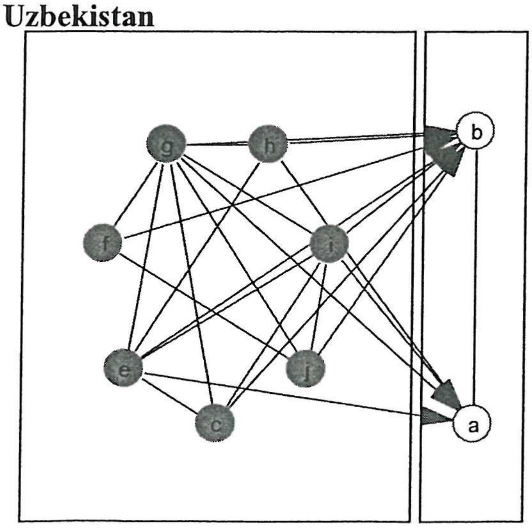
Key for all graphical chain models:

a	husband's age at marriage
b	wife's age at marriage
c	place of residence
d	married or cohabiting
e	household socio-economic status
f	wife works and earns cash
g	wife's highest level of education
h	husband's highest level of education
i	religion
j	ethnicity
l	household structure
m	consanguineous marriage
o	polygamous union

Peru



Uzbekistan



Bangladesh

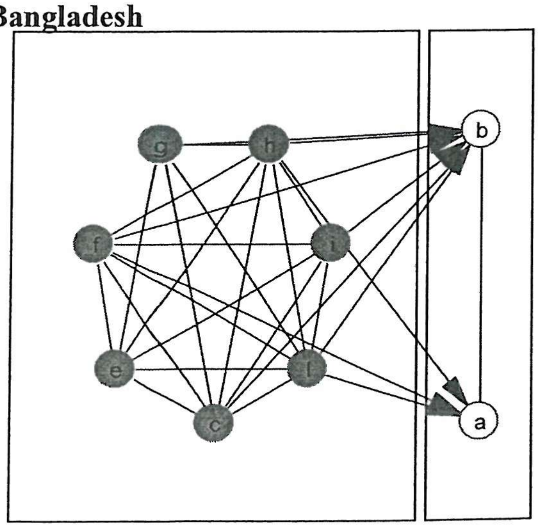
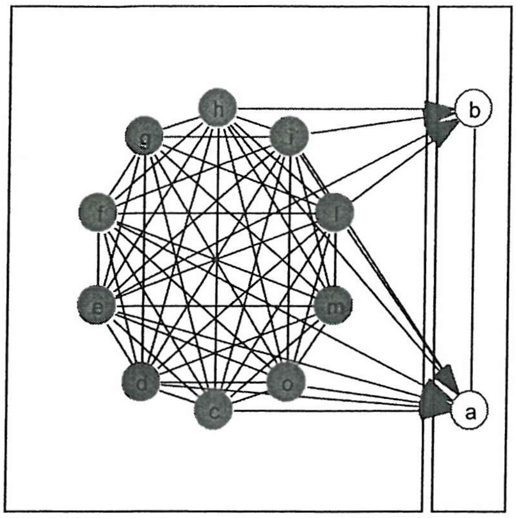
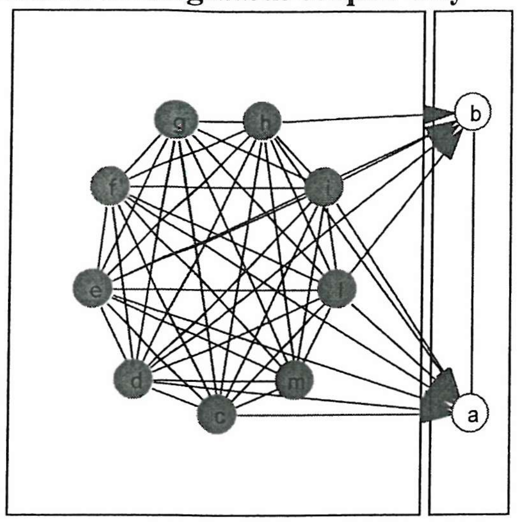


Figure 5.6 (continued) Parsimonious graphical chain models for five contrasting countries

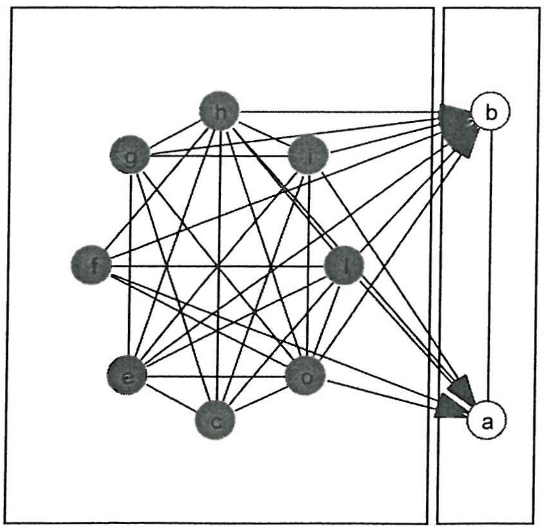
C.A.R.: All couples



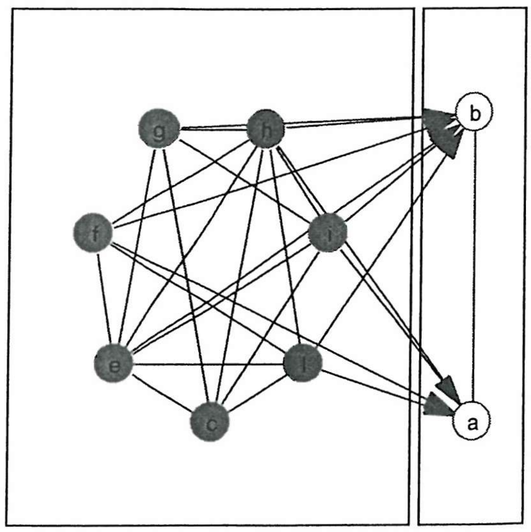
C.A.R.: Monogamous couples only



Burkina Faso: All couples



Burkina Faso: Monogamous couples only



While the statistical significance of the associations among the factors associated with age at marriage varies between countries, in all graphical chain models examined (including those not shown in the thesis), the edge remains between 'a' and 'b', i.e. between the ages at marriage of husbands and wives. The edge exclusion test is highly statistically significant (p=.000), which means that there remains a direct association between the ages at marriage of husbands and wives net of the marginal and conditional associations with the factors associated with age at marriage.

The association between the ages at marriage of husbands and wives can again be quantified with the partial correlation coefficient. However, in comparison to the partial correlation coefficients shown in Table 5.3, this is now net of the statistically significant marginal *and* conditional associations with the other variables. In Table 5.4, these 'new' partial correlation coefficients are labelled in bold as 'Partial (MIM)' as Table 5.4 also gives the partial correlation from using ordinary linear regression to model each spouse's age in turn, and the Pearson correlation coefficient, for reference.

Table 5.4 shows that, for all samples, there is little, if any difference between the partial correlation coefficient for the association between the ages at marriage of husbands and wives net of the marginal and conditional associations with the factors associated with age at marriage, and the partial correlation coefficient from regressing, in turn, the husband's age at marriage on these factors, and then the wife's age at marriage on these factors. As little difference was observed between the original partial correlation coefficient and the Pearson correlation coefficient (Table 5.4), there is thus little difference between the partial correlation coefficient obtained from the graphical modelling and the Pearson correlation coefficient. This can be interpreted as showing that there is little change in the magnitude of the association between the two ages at marriage net of the marginal and conditional associations with the factors associated with age at marriage.

As observed initially from the Pearson correlation coefficients for the 30 samples, there is much variation between samples in the extent of association between the ages at marriage of husbands and wives, net of the marginal and conditional associations with other variables. Figure 5.7 shows scatterplots of the corresponding partial residuals for husbands and wives, firstly for the Egyptian sample, and then for each of the five countries for which graphical chain models are presented above to illustrate this cross-national variation. The corresponding partial correlation coefficient from the graphical modelling is also given.

Table 5.4 Pearson and partial correlation coefficients as measures of the association between the ages at marriage of husbands and wives for all couples and for monogamous couples only, by country

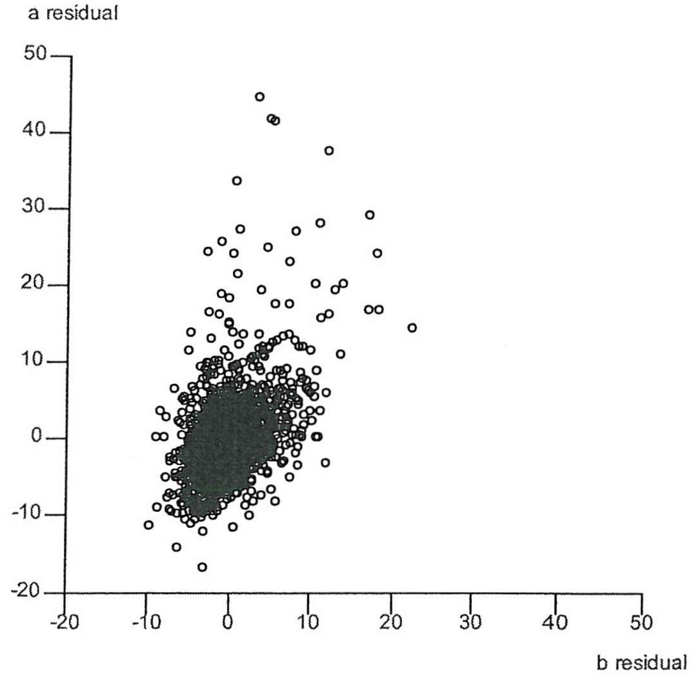
Sample:	All couples			Monogamous couples only		
Correlation coefficient:	Pearson	Partial [1]	**Partial [2] (MIM)**	Pearson	Partial [1]	**Partial [2] (MIM)**
Region and country [3]:						
Central & South America						
Guatemala	.45	.42	**.39**	.45	.42	**.39**
Peru	.56	.53	**.51**	.56	.53	**.51**
Brazil	.48	.48	**.44**	.48	.48	**.44**
Colombia	.50	.49	**.47**	.50	.49	**.47**
Dominican Republic	.40	.37	**.35**	.40	.37	**.35**
Haiti	.34	.39	**.33**	.34	.39	**.33**
West Asia & North Africa						
Kazakhstan	.61	.61	**.60**	.61	.61	**.60**
Uzbekistan	.67	.66	**.63**	.67	.66	**.63**
Egypt	.49	.43	**.43**	.49	.43	**.43**
South Asia						
Philippines	.48	.47	**.41**	.48	.47	**.41**
Nepal	.50	.47	**.47**	.52	.48	**.47**
Pakistan	.49	.49	**.47**	.52	.50	**.48**
Bangladesh	.38	.36	**.33**	.38	.36	**.33**
Central, East & South Africa						
Rwanda	.23	.21	**.20**	.26	.22	**.21**
Uganda	.43	.38	**.37**	.41	.36	**.34**
Malawi	.42	.41	**.41**	.47	.46	**.45**
Mozambique	.46	.45	**.45**	.52	.52	**.52**
Central African Republic	.39	.40	**.39**	.41	.41	**.40**
Zambia	.55	.51	**.51**	.55	.51	**.50**
Zimbabwe	.34	.39	**.38**	.47	.48	**.48**
Kenya	.35	.39	**.38**	.42	.43	**.43**
Tanzania	.37	.37	**.35**	.47	.47	**.46**
Comoros	.41	.35	**.33**	.41	.40	**.40**
West Africa						
Ghana	.43	.45	**.44**	.49	.48	**.46**
Benin	.26	.27	**.25**	.37	.30	**.29**
Niger	.26	.21	**.16**	.30	.23	**.18**
Burkina Faso	.08	.15	**.15**	.29	.26	**.21**
Cameroon	.24	.29	**.22**	.34	.32	**.32**
Côte d'Ivoire	.26	.33	**.34**	.42	.42	**.43**
Mali	.27	.27	**.24**	.37	.33	**.30**

Notes for Table 5.4

1. Net of the linear effects of the available factors associated with age at marriage –the marginal associations.
2. Net of the significant marginal *and* conditional associations between age at marriage and the available factors associated with age at marriage.
3. Countries are presented in order of magnitude of the observed median age difference (Table 4.1).

Figure 5.7 Scatterplot of the residuals from the hierarchical interaction model of the age at marriage of (a) husbands and (b) wives and factors hypothesised to be associated with age at marriage

Egypt: Partial correlation coefficient: .43



Peru: Partial correlation coefficient: .51

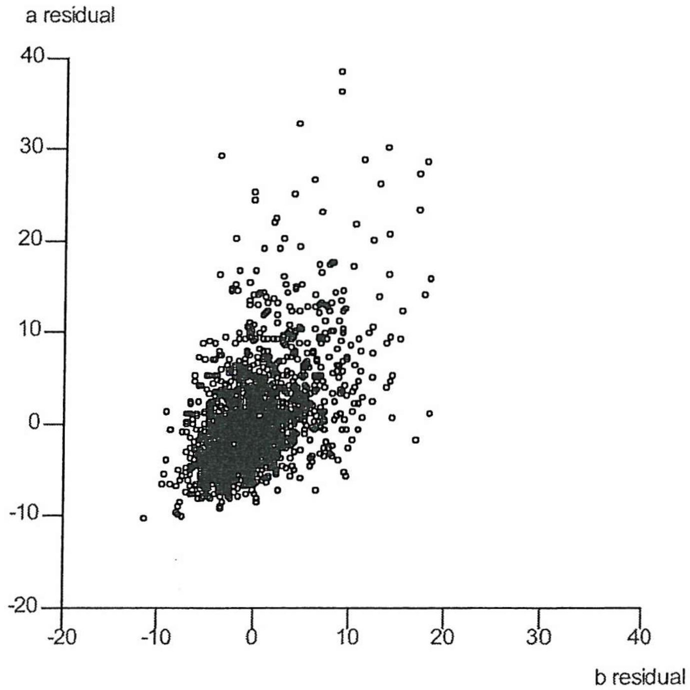
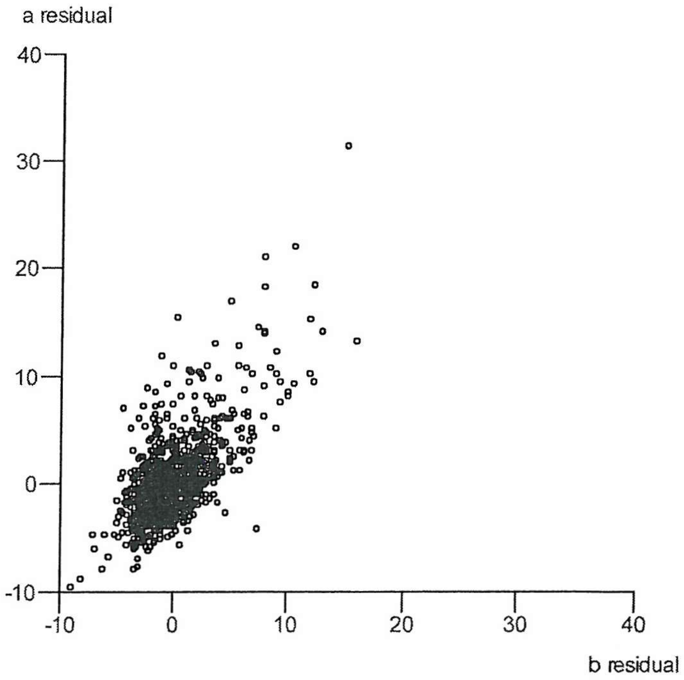


Figure 5.7 (continued) Scatterplot of the residuals from the hierarchical interaction model of the age at marriage of (a) husbands and (b) wives and factors hypothesised to be associated with age at marriage

Uzbekistan: Partial correlation coefficient: .63



Bangladesh: Partial correlation coefficient: .33

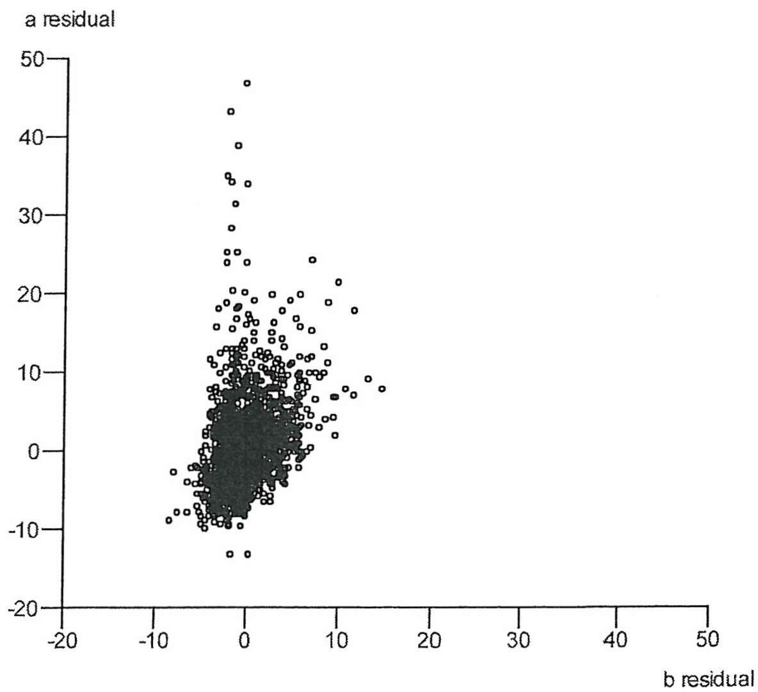
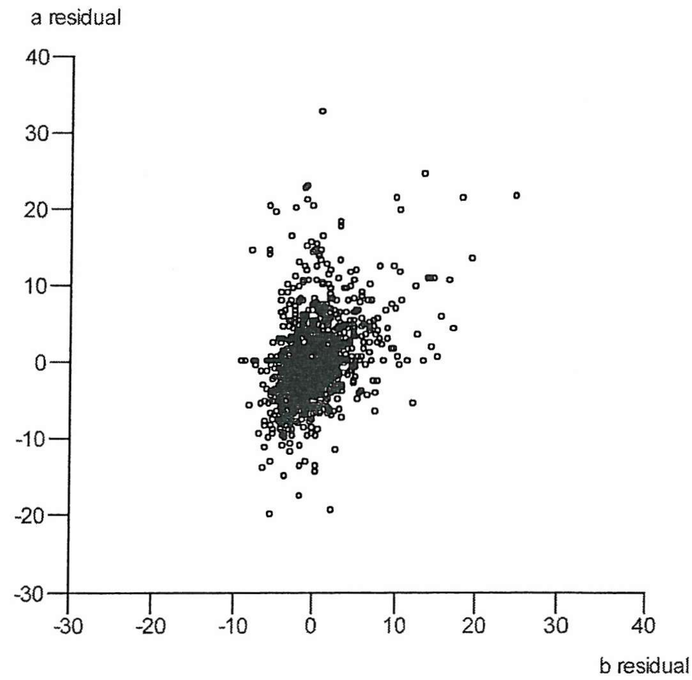


Figure 5.7 (continued) Scatterplot of the residuals from the hierarchical interaction model of the age at marriage of (a) husbands and (b) wives and factors hypothesised to be associated with age at marriage

C.A.R.: All couples: Partial correlation coefficient: .39



C.A.R.: Monogamous couples only: Partial correlation coefficient: .40

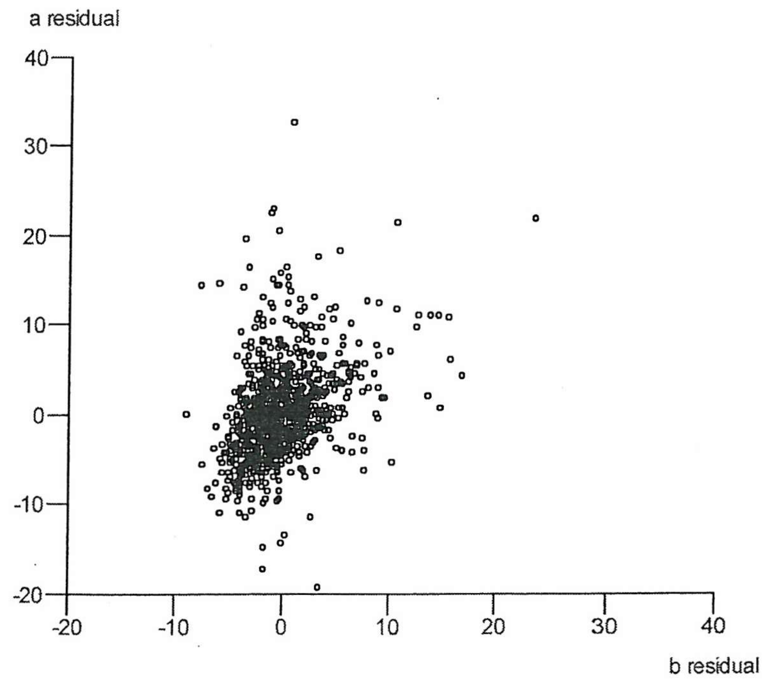
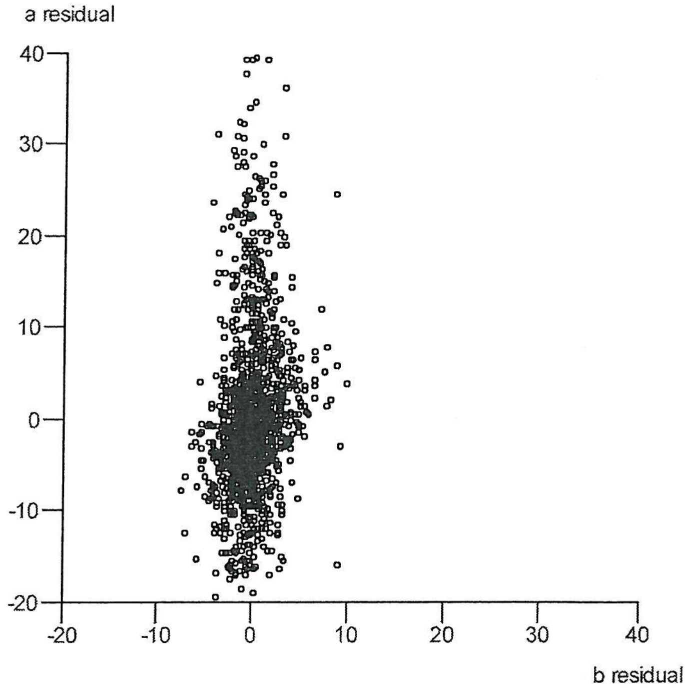
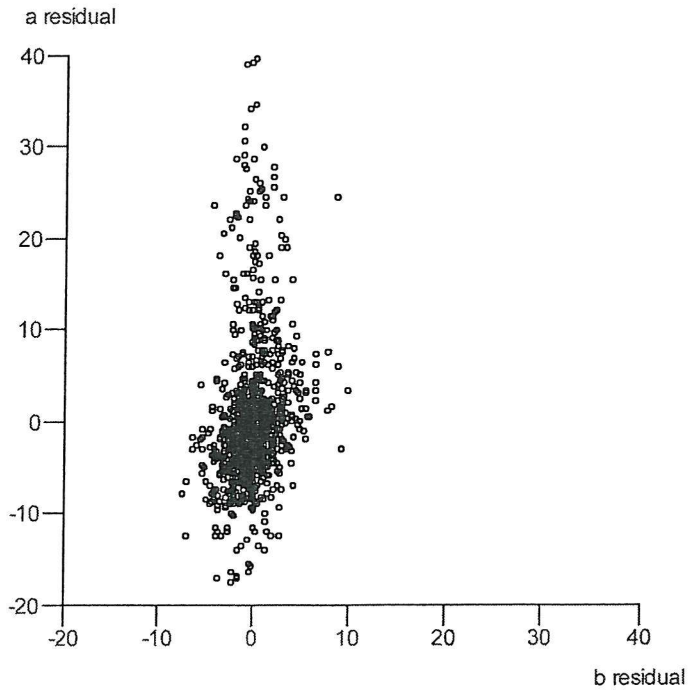


Figure 5.7 (continued) Scatterplot of the residuals from the hierarchical interaction model of the age at marriage of (a) husbands and (b) wives and factors hypothesised to be associated with age at marriage

Burkina Faso: All couples: Partial correlation coefficient: .15



Burkina Faso: Monogamous couples only: Partial correlation coefficient: .21



5.6 Graphical modelling the association between the ages at marriage of husbands and wives: Focussing on Egypt

The factors used in the hierarchical interaction models presented above have statistically significant direct associations with either the wife's age at marriage and/or the husband's age at marriage ($p<.05$). However, these factors do not account for much of the variation in either age at marriage when used as explanatory factors in a multiple regression of each spouse's age at marriage (not shown). For the Egyptian sample, less than 10% of the variation in the age at marriage of husbands is accounted for by these factors (adjusted R-square: 0.08), while approximately one-quarter of the variation in the age at marriage of wives is accounted for in this way (adjusted R-square: 0.27). One hypothesis could be that there are other factors that may have a greater association with age at marriage, perhaps to the extent that there is no direct association between the two ages at marriage, which would support the by-product hypothesis. The next section repeats the analysis presented in this chapter so far using data from the additional module on women's status collected as part of the 1995 Egyptian DHS survey, which includes factors specific to the time of marriage. This is a sub-sample of the Egyptian sample used so far since the women's status questionnaire was asked of a random sample of half the women eligible for the individual questionnaire. Thus, this analysis is based on 1102 couples (unweighted, 1168 weighted). Table 5.5 shows the variation in the median age at marriage (with the 95% confidence interval) for husbands and wives for each category of these factors and the letter codes used in the MIM analysis. For all factors, there is highly statistically significant variation in the age at marriage for husbands and wives ($p=.000$).

The five factors shown in Table 5.5 are coded: p, q, r, t and u, and are used with some of the factors used in the initial analysis for Egypt: household socio-economic status, wife's education, husband's education, and whether the union is consanguineous, codes e, g, h, and m, respectively (see Figure 5.5). Only some of the factors used in the initial analysis of the Egyptian data are used again here due to limitations regarding the size of the data matrix in MIM3.1. The religion factor, i, is excluded since it was not found to be directly associated with either the age at marriage of husbands or that of wives. In addition, some of the data available from this additional module can be considered as more informative since they refer to the time before the marriage rather than the time of the interview. For example, whether the wife lived in an urban or rural area at interview (code c in the previous analysis) has been superseded with the wife's childhood place of residence (code p in Table 5.5). Similarly, whether or not the wife worked for cash at the time of the interview (code f in the previous analysis) has been superseded with whether or not the wife worked before marriage (code q in Table 5.5). However, it is worth noting that there is a highly statistically significant association between factors c and p (p=.000 according to the Chi-square test), and also between factors f and q (p=.000 according to the Chi-square test). This suggests that for the other 29 samples where data referring to the time of marriage are not available, factors that refer to the time of interview may be sufficient proxies.

Table 5.5 Variations in the median ages at marriage of husbands and wives for factors hypothesised to be associated with age at marriage that are available from the women's status module of the 1995 Egyptian DHS

Factor, categories & MIM code	% of weighted sample (n=1169)	Age at marriage of: Husbands (95% CI)	Wives Median (95% CI)
Wife's childhood place of residence (p)		p=.000 [1]	p=.000 [2]
countryside	52.6	25.6 (25.2-26.1)	18.2 (17.9-18.4)
capital city/other city/town	47.3	27.8 (27.4-28.2)	21.5 (21.2-21.9)
	(100.0)		
Wife worked before marriage (q)		p=.000 [1]	p=.000 [2]
no	72.0	25.8 (25.5-26.2)	18.6 (18.4-18.9)
yes	28.0	28.7 (28.1-29.3)	22.7 (22.3-23.2)
	(100.0)		
Residence when first married (r)		p=.000 [1]	p=.000 [2]
with either husband's or wife's family	52.1	25.2 (24.8-25.6)	18.4 (18.1-18.7)
alone	47.9	28.3 (28.0-28.7)	21.3 (20.9-21.6)
	(100.0)		
Wife knew & chose husband (t)		p=.000 [3]	p=.000 [3]
had not met husband before marriage	41.8	27.8 (27.3-28.3)	20.2 (19.8-20.6)
knew husband but did not choose him herself	26.1	25.6 (25.1-25.7)	18.4 (18.0-18.8)
knew & chose husband herself	32.1	26.4 (26.0-26.8)	20.3 (19.9-20.7)
	(100.0)		
Husband's profession, grouped (u)		p=.000 [3]	p=.000 [3]
none/don't know	7.3	28.4 (26.9-29.9)	21.2 (20.1-22.3)
agriculture	20.1	24.4 (23.8-25.1)	17.4 (17.0-17.8)
skilled/unskilled production	27.9	26.3 (25.8-26.9)	19.3 (18.9-19.7)
professional/managerial/ technical/clerical/sales/service	44.7	27.8 (27.4-28.2)	21.1 (20.7-21.4)
	(100.0)		

Notes for Table 5.5

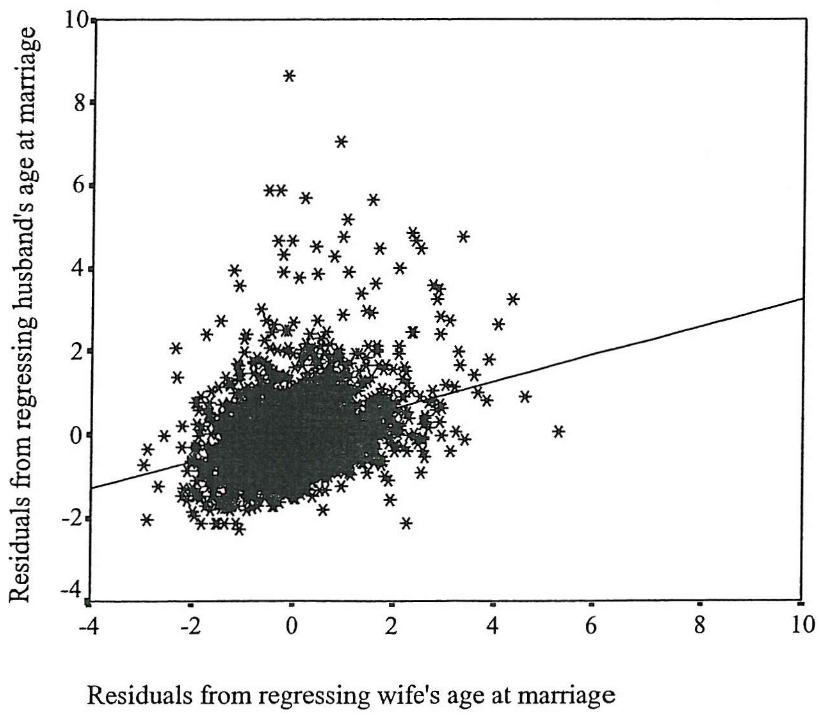
Statistical significant variation in the median age at marriage tested with:

1. t-test as variance homogeneity can be assumed
2. Mann-Whitney test as variance homogeneity can not be assumed
3. Kruskal-Wallis test as variance homogeneity can not be assumed

The partial correlation coefficient for the ages at marriage of husbands and wives net of the nine factors, e, g, h, m, p, q, r, t and u, is .32. This compares to the Pearson correlation coefficient of .40 for this reduced sample of eligible women who also completed the women's status module. This change in the correlation coefficient is similar in magnitude to the change from the Pearson correlation coefficient of .49 to the partial correlation coefficient of .43, as observed in Table 5.3 for the sample of all eligible women, net of seven factors c, e, f, g, h, i and m.

Figure 5.8 shows the association between the husband's age at marriage and the wife's age at marriage net of the linear effects of the nine factors for this sample of Egyptian couples. As before, this is achieved by plotting the residuals obtained from regressing the husband's age at marriage on these nine factors against the residuals obtained from regressing the wife's age at marriage on the same nine factors. The least-squares linear regression line that best fits the data is also displayed on the plot in Figure 5.8.

Figure 5.8 Scatterplot of the residuals from regressing each age at marriage on factors hypothesised to be associated with age at marriage, including some that are available from the women's status module of the 1995 Egyptian DHS

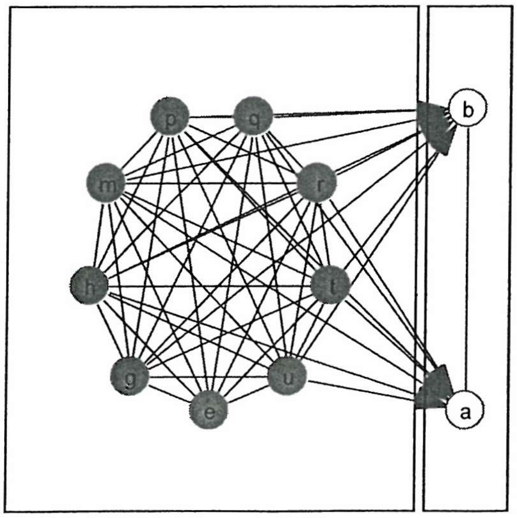


As observed above in Figure 5.3, the regression line in Figure 5.8 identifies a slight positive association between the residuals obtained from regressing the husband's age at marriage and the residuals obtained from regressing the wife's age at marriage. Thus, where the nine factors associated with age at marriage over-estimate the wife's age at marriage, it is likely that the same factors over-estimate the husband's age at marriage. However, as observed in the other previous residual plots (Figure 5.3), the association is far from perfect with much variation in the size of the residuals in both directions for the ages at marriage of both husbands and wives.

The joint distribution between the two ages at marriage net of the nine factors associated with age at marriage is now examined with a hierarchical interaction model using MIM. A parsimonious graphical model is identified from the stepwise selection procedure ($p<.05$) with the nine factors for this sample of Egyptian couples and is shown in Figure 5.9. It is can be expressed as:

1. eghmpqrtu
2. eghmpqrtu/bghmpqrtu,ahmpqrtu/ab

Figure 5.9 Graphical model of the associations between the ages at marriage of husbands and wives and factors associated with age at marriage, including some that are available from the women's status module of the 1995 Egyptian DHS



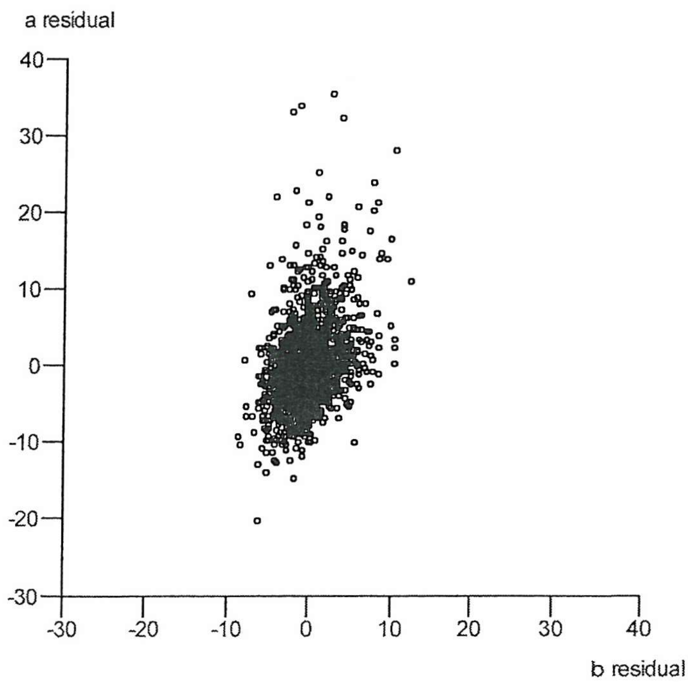
Key for Figure 5.9

a	husband's age at marriage	p	wife's childhood place of residence
b	wife's age at marriage	q	wife worked before marriage
e	household socio-economic status	r	post-marital residence when first married
g	wife's highest level of education	t	wife knew & chose husband
h	husband's highest level of education	u	husband's profession
m	consanguineous marriage		

Thus, in block 1 all nine factors that are associated with age at marriage are directly associated with each other. Net of the associations between the nine factors associated with age at marriage, all of the factors except household socio-economic status have a direct association with the wife's age at marriage, while all of the factors except household socio-economic status and wife's education have a direct association with the husband's age at marriage. Most importantly for this analysis is that again, there remains a direct association between the two ages at marriage net of marginal and conditional associations with the factors. The edge exclusion test identifies that this association is highly statistically significant (p=.000).

Having identified a parsimonious model, the corresponding partial residuals for husbands and wives are now plotted against each other in Figure 5.10. The corresponding partial correlation coefficient is .33. This compares to the Pearson correlation coefficient for this sample of .40, and a partial correlation coefficient obtained from the two separate linear regressions of .32. Thus, even after including some factors that may be considered as more specifically concerned with the timing of the marriage, there remains a positive net association between the ages at marriage of husbands and wives.

Figure 5.10 Scatterplot of the residuals from the hierarchical interaction model of the ages at marriage of husbands (a) and wives (b) and nine factors associated with age at marriage available from the women's status module of the 1995 Egyptian DHS



5.7 Discussion and conclusions

This chapter examined the hypothesis that the association between the ages at marriage of husbands and wives within-countries is a by-product of factors associated with age at marriage. Thus, this hypothesis assumed that the pattern of association between the husband's age and the wife's age is responsible for the systematic component of the distribution of age differences, i.e. the difference between the observed distribution and the distribution expected given the 'random' matching of spouses with respect to their ages at marriage. However, from the evidence presented in Chapter 5 it seems that the association is not just a by-product. This conclusion is reached from an analysis that took a lead from the analysis presented by Casterline *et al* (1986) but proceeded to use a more thorough approach. This analysis began by considering the partial correlation coefficient for the association between the ages at marriage of husbands and wives in each country, as Casterline *et al* (1986) reported. However, the analysis presented in this chapter considered this association net of as many as eleven factors that are hypothesised to be associated with age at marriage. In comparison, in their study of WFS data, Casterline *et al* (1986) just considered the association between the ages at marriage of husbands and wives net of two factors - the years of schooling achieved by the husband and the wife. Their justification for using these factors came from studies of assortative mating in Europe and North America, the relevance of which may be questioned given the traditional settings in which marriages occur in many of the less developed countries that they, and this chapter, studied (e.g. Fricke *et al*, 1986; Cain, 1993). Although, as discussed in Chapter 1, educational factors are increasingly relevant to the timing of marriage in more traditional settings not least because of the socio-economic implications (e.g. Jones, 1980; Lindenbaum, 1981; Zhang, 2000). Indeed, in most of the models derived for this chapter, a direct association was observed between the educational variables and the ages at marriage of husbands and wives. This suggests that Casterline *et al*'s (1986) use of just educational factors was perhaps justifiable.

A criticism of the method employed by Casterline *et al* (1986) is that it only considers marginal associations, which can result in identifying spurious associations (e.g. Edwards, 1995). A more thorough approach, and one which this chapter proceeded to adopt, is to use graphical modelling to consider the joint distribution of the ages at marriage of husbands and wives net of both marginal and conditional associations. From using these methods, the key finding of this chapter is that there is a statistically significant direct association between the husband's age at marriage and the wife's age at marriage for all 30 samples studied. Thus, despite the methodological improvements relative to the method used by Casterline *et al* (1986), and despite considering the association between the two ages at marriage net of as many as eleven factors, this chapter finds evidence in support of Casterline *et al*'s (1986, p. 362) conclusion that *"the association between the two ages of marriage is not simply due to assortative mating on determinants other than age"*. Indeed, it is worth recalling that this association not only remained but was considered highly statistically significant in all models (p=.000). This result applies to 30 culturally diverse countries, and remains regardless of the variables considered as factors associated with age at marriage that were available for each sample, the size of the sample, or whether the analysis was conducted on samples that included or excluded polygamous couples.

While the association between the ages at marriage of husbands and wives was considered net of a number of factors associated with age at marriage, it is possible that other, potentially more important factors may have been omitted from these analyses. Of course, this assumes that it is actually possible to identify omitted factors but, as noted in Chapter 1, it is difficult if not impossible to objectively conceptualise many of them, such as personality characteristics (Goldman *et al*, 1984). Indeed, it was difficult to find data in the DHS surveys that referred to the time of marriage for relatively objective factors, such as socio-economic status. Thus, it was necessary to use variables that referred to the time of interview as proxies for variables at the time of marriage for most of the analyses. The validity of this assumption could only be tested for a couple of variables from the Egyptian dataset as this is the only DHS survey of the 30 studied that collected data on characteristics at the time of marriage and at the time of interview. However, a highly statistically significant association was observed between the variables referring to the timing of marriage and the variables referring to the time of interview. None the less, using graphical models involved 'blocking' the variables, and thus the assumption that the variables in the first block occurred before the variables in the second block. This was not always strictly true because some of the variables referred to the time of interview. Consequently, it is important to recognise that direct associations identified by the models do not provide any indication of whether variables are positively or negatively related or the direction of the relationship. Thus, causality should not be assumed (e.g. Cox and Wermuth, 1996).

In light of these issues, it would have been useful to have used more data about the individual's circumstances at the time of marriage, such as those collected by the additional module on women's status as part of the 1995 Egyptian DHS. However, such data may be more likely to be missing, or at least, less reliable due to recall error (e.g. Singh, 1987; Chowdhury and Trovato, 1994). One possible solution may be to collect data from both the wife and her husband, which would allow internal consistency checks (as discussed in Chapter 2 with respect to age misreporting). Such an approach reiterates the demand for nuptiality studies to adopt a *"two-sex perspective"* (e.g. Schoen, 1983; Xenos and Gultiano, 1992). However, at the time of writing, only eleven DHS surveys have included a questionnaire for husbands since the DHS programme began in 1984.

In conclusion, while the data and methods used for this analysis are not perfect, the approach adopted is an improvement on that used by Casterline *et al* (1986). Although, despite these efforts, this chapter reaches the same conclusion as they reported. Thus, the association between the ages at marriage of husbands and wives is not simply a by-product of factors associated with age at marriage. While Chapter 4 reported evidence in support of the availability hypothesis, the preferences hypothesis may also apply. Since the DHS surveys do not collect data on age preferences, as previously noted, Chapter 6 proceeds to examine an associated hypothesis, that the age difference reflects women's status at the individual level.

Chapter 6:
Results of examining the hypothesis that the age difference reflects women's status at the individual level in Egypt

6.1 Overview

This chapter examines the women's status hypothesis at the individual level in the context of Egypt since, at the time of analysis, this was the only DHS survey to include questions on women's status, or rather women's autonomy, within the household. Based on a method developed by Kishor (1995) who analysed similar data from the 1988 Egyptian DHS, responses to 15 questions are scored, and from these, three autonomy indices are derived. Variations in the average age difference are firstly considered by response to each of the questions used to create the autonomy indices. Variations in the average age difference are then considered by each level of each autonomy index. The analysis then uses graphical modelling to consider the association between these indices and the age difference net of the association with the ages at marriage of husbands and wives. This is to examine the hypothesis that an association between the age difference and women's autonomy is due, in a statistical sense at least, to the association between the age difference and the wife's age at marriage, and the association between the wife's age at marriage and women's autonomy. For completion, the chapter concludes by using graphical modelling to examine the association between the ages at marriage of husbands and wives net of the autonomy indices and the factors identified in Chapter 5 as statistically significantly associated with age at marriage.

6.2 Variations in the distributions of the indicators of women's autonomy

Table 6.1 shows the percentage of women reporting either the less autonomous or the more autonomous responses for each of the 15 variables used to derive the three autonomy indices. The question for which the largest proportion of women in this sample report the more autonomous response is, *"who has the last word on the use of family planning methods?"* 89.9% of women reported that they made decisions regarding the use of contraception either solely or jointly with their husband. In contrast, the question for which the smallest proportion of women report the more autonomous response is, *"is the wife allowed to go alone to the* *neighbourhood for recreation?"* 13.3% of women reported that were allowed to go alone to the neighbourhood for recreation.

Of the three sets of variables, the customary autonomy variables have the highest mean percentage of women reporting a more autonomous responses at 79.8%, which compares to means of 59.2% and 50.3% for the non-customary and realised autonomy variables, respectively. As might be anticipated, there is some association between the responses reported for each of the variables used to define each index. 743 of the 1168 women in this weighted sample (63.7%) reported the same response for all five questions used to derive the customary index, i.e. always the less autonomous response or always the more autonomous response. Most of these 743 women consistently reported the more autonomous response (90.3% of the 743 women). In contrast, 193 of the 1168 women (16.5%) reported the same response for all five questions used to derive the non-customary autonomy index. Most of these 193 women consistently reported the more autonomous response (82.4% of the 193 women). For the realised autonomy index, 194 of the 1168 women (16.6%) reported the same response for all five variables used to derive this index. Of these 194 women, approximately one-third consistently reported the more autonomous response (36.6% of the 194 women). Overall, only seven of the 1168 women (0.6%) reported the less autonomous response for all 15 questions, and 24 of the 1168 women (2.1%) reported the more autonomous response for all 15 questions.

These associations can be quantified with Cronbach's Alpha, a measure of each index's internal reliability, which is calculated as 0.815 for the customary autonomy index, 0.418 for the non-customary autonomy index, and 0.576 for the realised autonomy index. As Cronbach's Alpha is greatest for the customary autonomy index, this suggests that the variables are most alike for this index. In contrast, the variables used to derive the non-customary autonomy index are the least alike as Cronbach's Alpha is the smallest for this index.

Table 6.1 Percentage distribution of responses to 15 questions used to derive three women's autonomy indices

Questions by index	% women reporting [2]		
	Less autonomous response [3]	**More autonomous response** [3]	**Total**
Customary autonomy index [1]			
Who has the last word on...?	Husband/other:	Wife/ both husband & wife:	
Having another child	22.7	77.3	100.0
Children's education	22.8	77.2	100.0
Children's marriage plans	27.3	72.7	100.0
Medical attention for the children	18.1	81.9	100.0
Use of family planning methods	10.1	89.9	100.0
Non-customary autonomy index [1]			
Who has the last word on...?	Husband/other:	Wife/ both husband & wife:	
Visits to friends and family	58.2	41.8	100.0
Household budget	52.9	47.1	100.0
Agree/disagree...	Disagree/ don't know:	Agree:	
A wife's place is not only in the home	27.0	73.0	100.0
If a wife disagrees with her husband, she should speak-up	38.6	61.4	100.0
Wife knows her family's income (approximately)	No: 27.1	Yes: 72.9	100.0
Realised autonomy index [1]			
Is the wife allowed to go alone to...	No:	Yes:	
Local market to buy things	20.0	80.0	100.0
Local health centre/doctors	34.9	65.1	100.0
Neighbourhood for recreation	86.7	13.3	100.0
Home of friends/relatives	51.9	48.1	100.0
Who mainly decides how family income is spent?	Husband/other: 55.1	Wife/ both husband & wife: 44.9	100.0

Notes for Table 6.1

1. These questions are similar to the questions used to derive three indices of women's autonomy based on a method reported by Kishor, S. (1995) *Autonomy and Egyptian women: Findings from the 1988 Egypt Demographic and Health Survey*. Demographic and Health Surveys Occasional Paper no. 2. Calverton, Maryland: Macro International Inc. See Appendix 2L for differences between the questions used for this analysis and the questions used by Kishor (1995).
2. Weighted base of 1168 women.
3. See Table 2.5 for a detailed breakdown as to which responses are considered as more vs. less autonomous.

6.3 Variations in the median age difference by response to women's autonomy questions

Table 6.2 shows the median age difference and its 95% confidence interval for each type of response to each of the 15 questions used to derive the three autonomy indices. Questions for which there is statistically significant variation in the age difference by response are highlighted in bold. Statistically significant variation in the average age difference is observed for only two of the 15 variables used to derive the indices of women's autonomy ($p<.05$). Women who solely or jointly have the last word regarding medical attention for their children have a statistically significantly smaller age difference, on average, than women whose husband or someone else has the last word, with medians of 5.5 years and 6.0 years respectively ($p=.008$). This association is in the hypothesised direction, i.e. women reporting a more autonomous response tend to have smaller age differences, on average. In contrast, women who agreed that they should speak up if they disagreed with their husband have a statistically significantly larger age difference, on average, than women who disagreed with this statement, with medians of 5.9 years and 5.1 years respectively ($p=.045$).

However, given that fifteen variables are analysed here, statistically significant variation is expected to occur by chance for one or possibly two variables assuming $p<.05$. Thus, it is reasonable to conclude from this analysis that there is little evidence in support of the hypothesis that at the individual level the age difference reflects women's status, at least as measured by these hypothesised indicators of women's autonomy.

Table 6.2 Variations in the median age difference by response to 15 questions used to derive three women's autonomy indices

Questions by index	Median age difference (and 95% C.I.) by response to women's autonomy questions: Less autonomous response [2]	More autonomous response [2]	*p*-value [3]
Customary autonomy index [1]			
Who has the last word on...?	Husband/other:	Wife/ both husband & wife:	
Having another child	5.8 (5.1-6.6)	5.6 (5.2-6.0)	.077 [5]
Children's education	5.6 (4.9-6.4)	5.6 (5.2-6.0)	.107 [5]
Children's marriage plans	5.5 (5.0-6.1)	5.7 (5.2-6.1)	.481 [5]
Medical attention for the children	**6.0 (5.1-6.9)**	**5.5 (5.2-5.9)**	**.008** [4]
Use of family planning methods	5.1 (4.1-6.5)	5.7 (5.3-6.0)	.281 [4]
Non-customary autonomy index [1]			
Who has the last word on...?	Husband/other:	Wife/ both husband & wife	
Visits to friends and family	5.4 (4.9-5.9)	5.8 (5.4-6.2)	.439 [4]
Household budget	5.7 (5.0-6.3)	5.6 (5.2-5.9)	.478 [5]
Agree/disagree...	Disagree/ don't know:	Agree:	
A wife's place is not only in the home	5.2 (4.6-5.9)	5.8 (5.4-6.2)	.376 [5]
If a wife disagrees with her husband, she should speak-up	**5.1 (4.6-5.7)**	**5.9 (5.5-6.3)**	**.045** [5]
	No:	Yes:	
Wife knows her family's income (approximately)	5.6 (4.9-6.3)	5.6 (5.2-6.0)	.473 [5]
Realised autonomy index [1]			
Is the wife allowed to go alone to...	No:	Yes:	
Local market to buy things	5.5 (4.6-6.4)	5.6 (5.3-6.0)	.535 [5]
Local health centre/doctors	5.6 (5.1-6.1)	5.6 (5.2-6.0)	.680 [5]
Neighbourhood for recreation	5.5 (5.1-5.8)	6.5 (5.6-7.3)	.087 [5]
Home of friends/relatives	5.4 (4.9-6.0)	5.8 (5.4-6.2)	.345 [5]
Who mainly decides how family income is spent?	Husband/other: 5.7 (5.1-6.2)	Wife/ both husband & wife: 5.6 (5.2-5.9)	.161 [4]

Notes for Table 6.2

1. See note 1 for Table 6.1.
2. See note 3 for Table 6.1.
3. Statistically significant variations in the average age difference are highlighted in bold ($p<.05$).
4. Statistical significance tested with the Mann-Whitney test since variance homogeneity can not be assumed.
5. Statistical significance tested with the t-test since variance homogeneity can be assumed.

6.4 Deriving three women's autonomy indices

Three indices of women's autonomy are derived according to a method used by Kishor (1995, see section 2.4.2). The less autonomous responses to the questions shown in Tables 6.1 and 6.2 each score zero and the more autonomous responses shown in these tables each score one. The responses to each group of five questions are then summed to make three autonomy indices for each woman. For all women, the minimum score for each index is zero and the maximum score is five. Figure 6.1 shows the percentage distribution of the summed scores for each index.

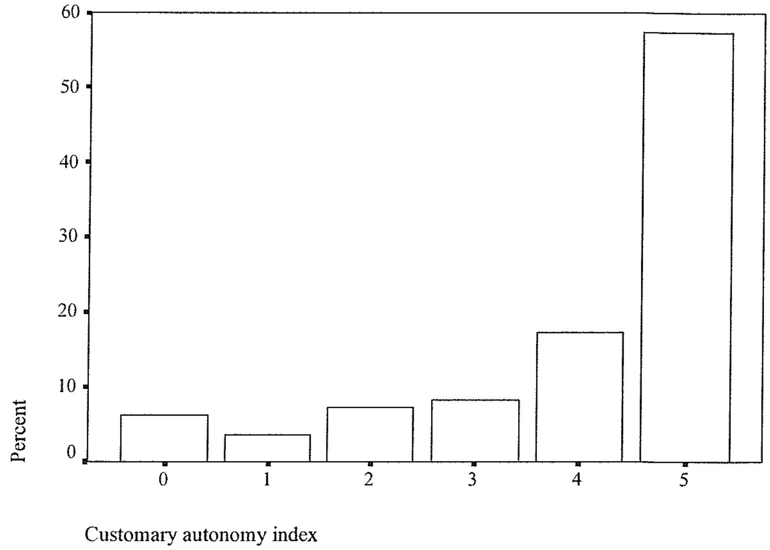
The distribution of scores for the customary autonomy index is negatively skewed with over half of the women in this sample (57.4%) scoring the maximum for this index, i.e. consistently reporting the more autonomous response, and thus having the highest level of customary autonomy according to this measure. For this index, the remaining scores consist of, at most, 17.2% of the sample, with just 6.2% of women scoring zero. In contrast, the indices of non-customary autonomy and realised autonomy are approximately Normally distributed. For both of these indices, three is the modal score, which constitutes approximately one-quarter of each sample (27.9% and 28.8% respectively).

The scores are then categorised for each index to create three levels of autonomy: low, medium and high. For the non-customary autonomy and realised autonomy indices, women scoring 0 or 1 are considered as having low autonomy; women scoring 2 or 3 are considered as having medium autonomy, and women scoring 4 or 5 are considered as having high autonomy. However, given the skewed nature of the distribution of customary autonomy scores, women scoring 0,1, or 2 are considered as having low autonomy; women scoring 3 or 4 are considered as having medium autonomy, and women scoring the maximum, 5, are considered as having high autonomy.

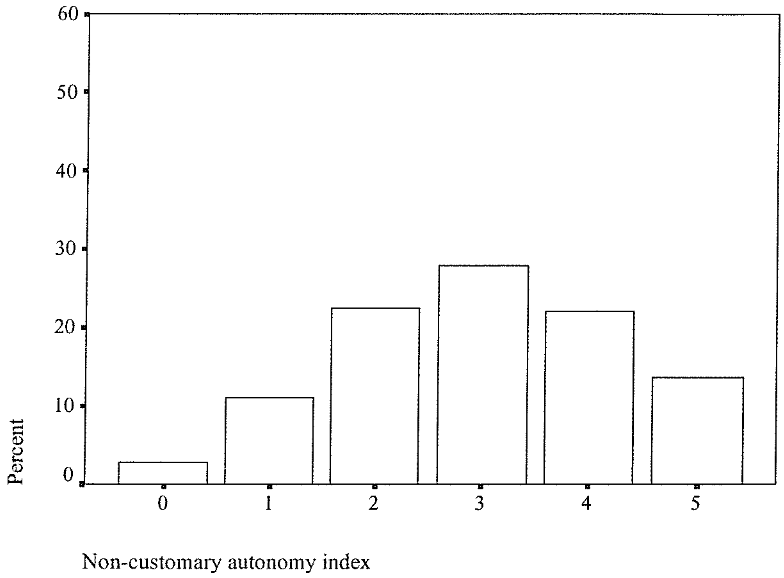
Table 6.3 shows the cross-tabulation of the three autonomy indices. All three indices are highly statistically significantly associated (p=.000 according to the Chi-square statistic). Approximately one-quarter of the 1168 women have the same level of autonomy for all three indices (25.2%). 3.5% of all women have low autonomy according to all three indices; 9.6% of all women have medium autonomy according to all three indices; and 12.1% of all women have high autonomy according to all three indices. In contrast, 10.1% of all women have a different level of autonomy for each index.

Figure 6.1 Percentage distribution for each women's autonomy index

Customary autonomy index



Non-customary autonomy index



Realised autonomy index

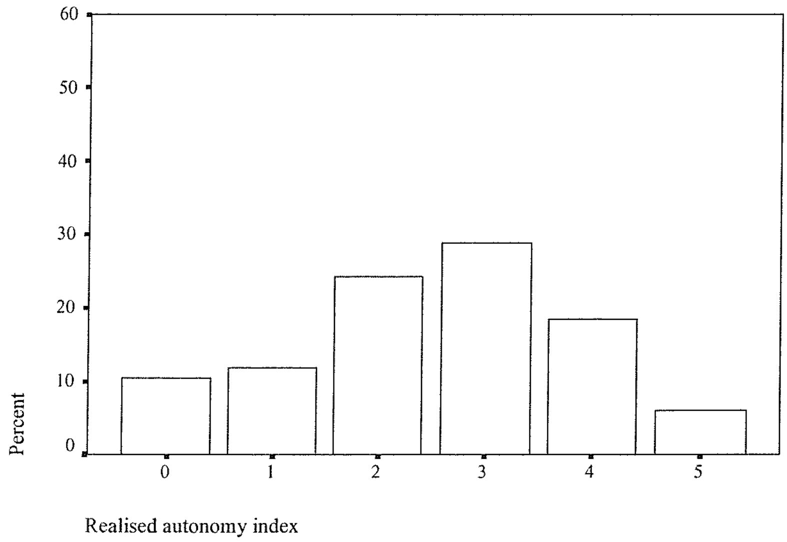


Table 6.3 Cross-tabulation of three indices of women's autonomy

Level of realised autonomy index [1]	Level of non-customary autonomy index [1]: Low	Medium	High	(Total [2])
Low				
Level of customary autonomy index [1]				
Low	41	38	3	(82)
Row %	*50.0*	*46.3*	*3.7*	*100.0*
Column %	*55.4*	*24.5*	*9.7*	*31.5*
Medium	14	54	6	(74)
Row %	*18.9*	*73.0*	*8.1*	*100.0*
Column %	*18.9*	*34.8*	*19.4*	*28.5*
High	19	63	22	(104)
Row %	*18.3*	*60.6*	*21.2*	*100.0*
Column %	*25.7*	*40.6*	*71.0*	*40.0*
(Total)	(74)	(155)	(31)	(260)
(Row %)	*(28.5)*	*(59.6)*	*(11.9)*	*(100.0)*
(Column %)	*(100.0)*	*(100.0)*	*(100.0)*	*(100.0)*
Medium				
Level of customary autonomy index [1]				
Low	28	51	7	(86)
Row %	*32.6*	*59.3*	*8.1*	*100.0*
Column %	*38.4*	*15.5*	*3.2*	*13.9*
Medium	19	112	46	(177)
Row %	*10.7*	*63.3*	*26.0*	*100.0*
Column %	*26.0*	*34.1*	*21.1*	*28.6*
High	26	165	165	(356)
Row %	*7.3*	*46.3*	*46.3*	*100.0*
Column %	*35.6*	*50.3*	*75.7*	*57.5*
(Total)	(73)	(328)	(218)	(619)
Row %	*(11.8)*	*(53.0)*	*(35.2)*	*(100)*
Column %	*(100.0)*	*(100.0)*	*(100.0)*	*(100.0)*
High				
Level of customary autonomy index [1]				
Low	9	16	6	(30)
Row %	*26.7*	*53.3*	*20.0*	*100.0*
Column %	*53.3*	*15.2*	*3.6*	*10.4*
Medium	0	26	21	(47)
Row %	*0.0*	*55.3*	*44.7*	*100.0*
Column %	*0.0*	*24.8*	*12.5*	*16.3*
High	7	63	141	(211)
Row %	*3.3*	*29.9*	*66.8*	*100.0*
Column %	*46.7*	*60.0*	*83.9*	*73.3*
(Total)	(16)	(105)	(168)	(289)
Row %	*5.2*	*36.5*	*58.3*	*(100.0)*
Column %	*(100.0)*	*(100.0)*	*(100.0)*	*(100.0)*

Notes for Table 6.3
1. See note 1 for Table 6.1.
2. Weighted base of 1168 women

6.5 Variations in the median age difference by level of women's autonomy

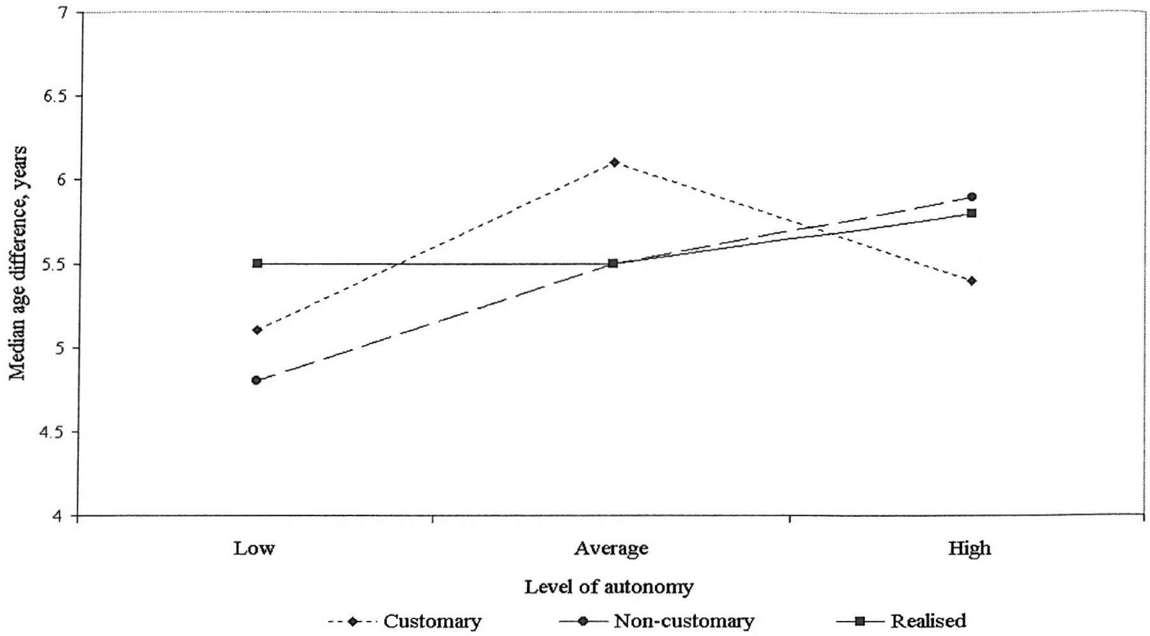
Table 6.4 shows the median age difference and its 95% confidence interval for each level of each autonomy index. Figure 6.2 displays these data graphically. The customary autonomy index is the only index with statistically significant variation in the average age difference (p=.025 according to the Kruskal-Wallis test). The median age difference for women with low autonomy according to the customary autonomy index is 5.1 years, while for women with medium autonomy according to this index, the median age difference is 6.1 years, which is smaller than the median for women with high autonomy at 5.4 years. Indeed, there is no statistically significant variation in the average age difference for women with low autonomy relative to women with medium autonomy, or between women with low autonomy relative to women with high autonomy, according to Mann-Whitney tests. However, age differences are statistically significantly smaller for women with high autonomy relative to women with medium autonomy (p=.018). Thus, the hypothesis that age differences tend to be smaller on average among women with higher autonomy is only partly supported with these data. While not statistically significant, the median age difference is smallest for women with low autonomy for both the non-customary autonomy and realised autonomy indices, which contradicts this hypothesis.

Table 6.4 Median age difference for each level of autonomy and each autonomy index

Autonomy index [1]	Level of women's autonomy			*p*-value
	Low	**Medium**	**High**	
customary (weighted n)	5.1 (4.4-6.2) (199)	6.1 (5.6-6.7) (298)	5.4 (4.9-5.9) (671)	.025 [2]
non-customary (weighted n)	4.8 (4.3-5.8) (162)	5.5 (4.9-6.1) (589)	5.9 (5.5-6.4) (417)	.743 [2]
realised (weighted n)	5.5 (4.8-6.4) (261)	5.5 (5.1-6.0) (619)	5.8 (5.2-6.6) (288)	.748 [3]

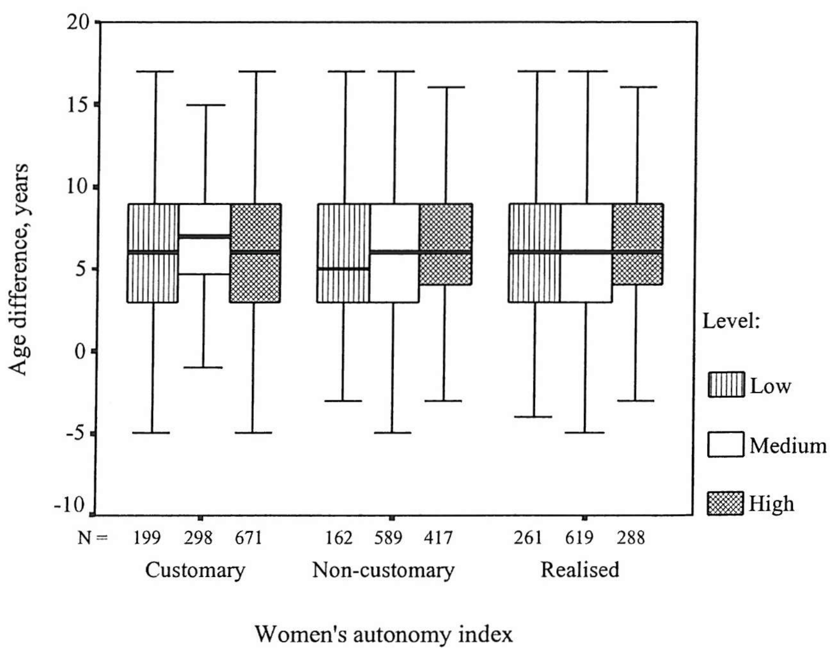
Notes for Table 6.4
1. See note 1 for Table 6.1
2. Statistical significance tested using the Kruskal-Wallis test since variance homogeneity can not be assumed.
3. Statistical significance tested using the analysis of variance since variance homogeneity can be assumed.

Figure 6.2 Line graph showing variations in the median age difference by level of autonomy for three autonomy indices



Despite some statistically significant variation in the average age difference, as Figure 6.3 shows, there is very little difference in the distribution of the age difference by level of autonomy for each autonomy index.

Figure 6.3 Boxplots to show the extent of variation in the distributions of age differences by level of autonomy and women's autonomy index



6.6 Associations between the age difference, age at marriage and autonomy

As observed in Chapter 4, there is a greater association within countries between the age difference and the age at which men marry, than the age at which women marry. Given the associations between (1) the age difference and age at marriage; and (2) the age difference and some of the hypothesised indicators of women's autonomy, variations in the ages at marriage of husbands and wives are now considered by level of autonomy for each autonomy index. Table 6.5 shows the median ages at marriage of wives and husbands (and 95% confidence intervals) for each level of autonomy for each autonomy index. Figure 6.4 displays these medians graphically.

Table 6.5 The median age at marriage of wives and husbands, respectively, by level of autonomy for each autonomy index

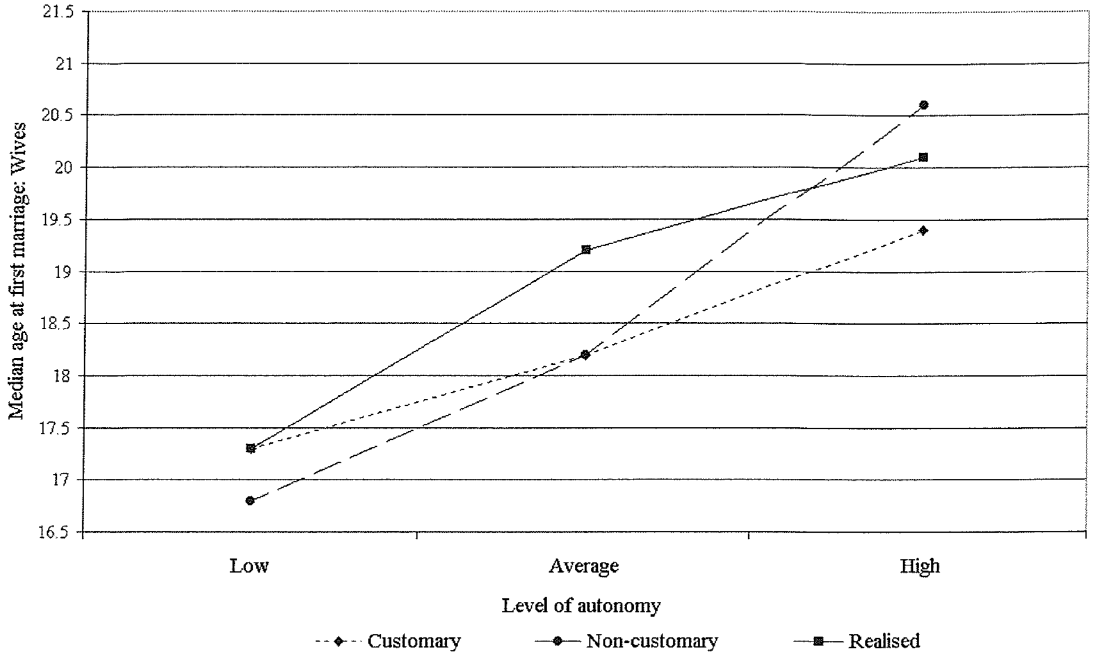
Autonomy index [1]	Level of autonomy [1]			*p*-value
	Low	Medium	High	
Wives				
Customary [1]	17.3 (16.8-18.1)	18.2 (17.5-19.2)	19.4 (19.0-19.9)	.000 [2]
Non-customary [1]	16.8 (16.2-17.5)	18.2 (17.8-18.7)	20.6 (20.3-20.9)	.000 [3]
Realised [1]	17.3 (16.8-17.9)	19.2 (18.7-19.7)	20.2 (19.9-20.6)	.000 [3]
Husbands				
Customary [1]	24.9 (24.0-25.7)	25.9 (25.4-26.7)	26.2 (25.8-26.6)	.003 [3]
Non-customary [1]	23.4 (22.5-24.4)	25.2 (24.8-25.6)	27.4 (26.9-27.9)	.000 [3]
Realised [1]	24.3 (23.5-25.2)	25.8 (25.5-26.3)	26.8 (26.4-27.3)	.000 [2]

Notes for Table 6.5

1. See note 1 for Table 6.1.
2. Statistical significance tested using an analysis of variance as variance homogeneity can be assumed.
3. Statistical significance tested using Kruskal-Wallis test as variance homogeneity can not be assumed.

Figure 6.4 Line graphs showing variations in the median age at marriage for wives and husbands respectively by level of autonomy for each autonomy index

Wives



Husbands

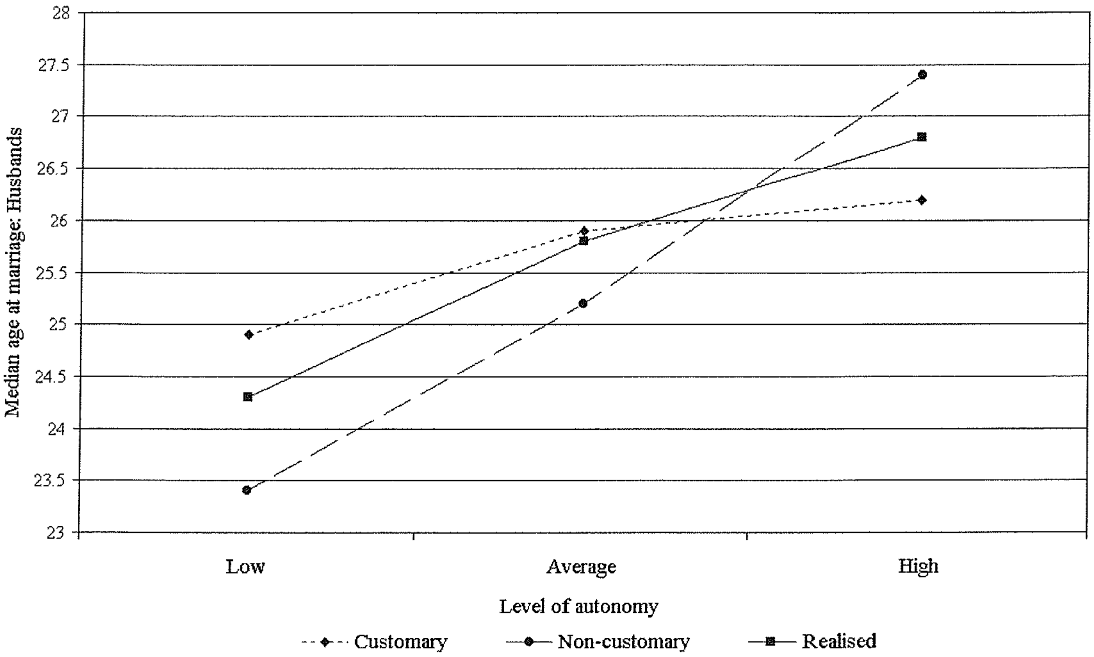


Table 6.5 shows for each autonomy index, there is highly statistically significant variation in the ages at marriage of both husbands and wives ($p<.01$). As Figure 6.4 shows, there is a positive association such that the average age at marriage -for husbands and wives - increases with increasing level of women's autonomy. The greatest variation in the median age at marriage is observed for the non-customary autonomy index for both wives and husbands. There is 3.8 years difference between the median ages at marriage of wives with low non-customary autonomy and wives with high non-customary autonomy. Similarly for husbands, there is four years difference between the median ages at marriage of husbands whose wives have low non-customary autonomy and husbands whose wives have high non-customary autonomy.

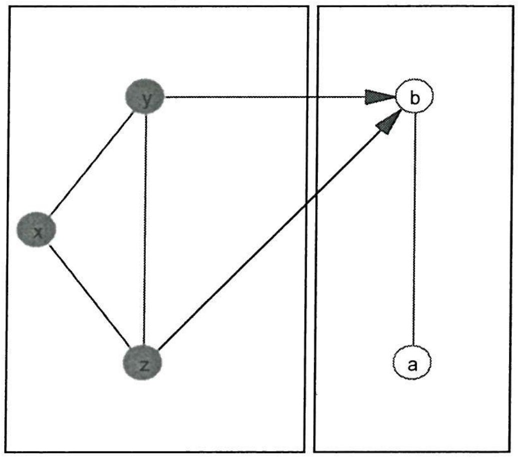
These results are interesting given that there is a highly statistically significant association between the age at marriage of husbands and the age difference within countries, but that there is little statistically significant variation in the age difference by level of autonomy. However, it is important to recall the association between the ages at marriage of husbands and wives at this point. As Chapter 5 identified, there is a highly statistically significant association between the ages at which spouses marry, which is not a by-product of factors associated with age at marriage. It is possible therefore that the statistically significant variation in the husband's age at marriage (Table 6.5) and the age difference (Table 6.4, albeit just for customary autonomy) by level of autonomy is spurious, due to the statistically significant association between the husband's age at marriage and the wife's age at marriage, and the statistically significant association between the autonomy indicators and the wife's age at marriage. Since the age difference net of the wife's age at marriage as an outcome is statistically equivalent to the husband's age at marriage and *vice versa* (see section 2.2.6), graphical modelling is used in the next section to examine the association between this 'net' age difference and the women's autonomy indices.

6.7 Associations between the 'net' age difference and women's autonomy indices

As in Chapter 5, the graphical modelling package MIM3.1 is used to derive a hierarchical interaction model of the association between the women's autonomy indices, the husband's age at marriage and the wife's age at marriage. In MIM, the autonomy indices are labelled 'x' for the customary autonomy index, 'y' for the non-customary autonomy index, and 'z' for the realised autonomy index. As in Chapter 5, the husband's age at marriage is labelled 'a' and the wife's age at marriage is labelled 'b'. The autonomy indices are in block one, and the ages at marriage are in block two, as in Chapter 5. A parsimonious graphical model is identified from the stepwise selection procedure in relation to these five variables and is shown in Figure 6.5. It is expressed as:

1. xyz
2. xyz/byz, a/ab

Figure 6.5 A graphical model of the associations between the ages at marriage of husbands and wives and three indices of women's autonomy derived from 15 variables from the women's status module of the 1995 Egyptian DHS using a method developed by Kishor (1995)



Key for Figure 6.5

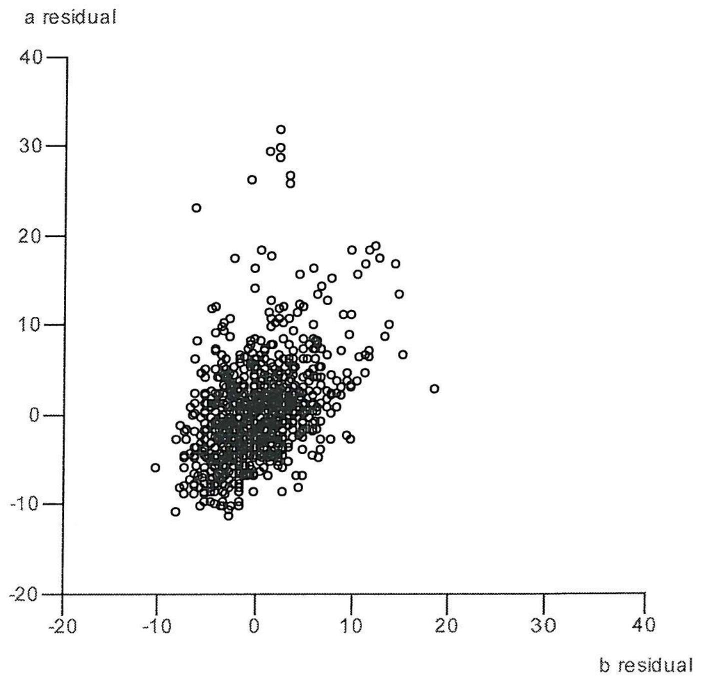
a	husband's age at marriage
b	wife's age at marriage
x	customary autonomy index
y	non-customary autonomy index
z	realised autonomy index

As observed from the three-way cross-tabulation of the autonomy indices in Table 6.3, the three autonomy indices are all statistically significantly associated with each other in the graphical model ($p<.05$). In Figure 6.5, the association between the customary autonomy and realised autonomy indices is statistically significant at $p=.016$, while the other associations are significant at $p=.000$. Net of the associations in block one and the association between the ages at marriage of husbands and wives in block two, the non-customary autonomy and realised autonomy indices have highly statistically significant direct associations with b, i.e.

the age difference net of the husband's age at marriage (p<.001). However, net of the associations in blocks 1 and 2, none of the autonomy indices have a statistically significant direct association with a, i.e. the age difference net of the wife's age at marriage. This result supports the hypothesis that the individual-level association between the age difference and women's autonomy is due, in a statistical sense at least, to the association between the age difference and the wife's age at marriage, and the association between the wife's age at marriage and women's autonomy.

It is also worth noting that, as observed throughout the examination of the by-product hypothesis in Chapter 5, there remains a direct association between the ages at marriage of husbands and wives net of the marginal and conditional associations with the autonomy indices. The edge exclusion test identifies that this association is highly statistically significant (p=.000). The partial residuals for husbands and wives are plotted against each other in Figure 6.6. The corresponding partial correlation coefficient is .47, which compares to the Pearson correlation coefficient for this sample of .52.

Figure 6.6 Scatterplot of the residuals from the hierarchical interaction model of the ages at marriage of husbands (a) and wives (b) and three women's autonomy indices



6.8 Associations between the ages at marriage of husbands and wives, the women's autonomy indices and factors associated with age at marriage

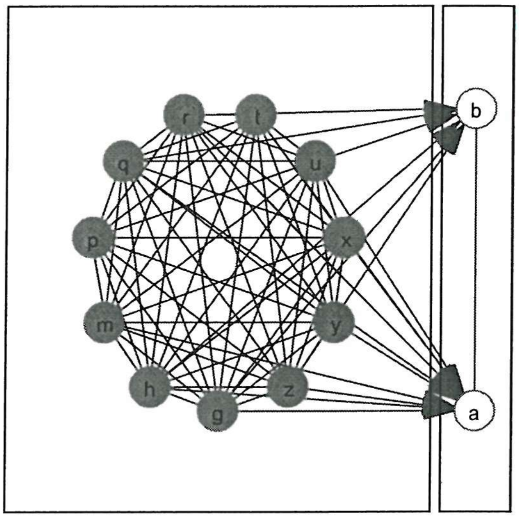
For completion, the final section of analyses examines the association between the ages at marriage of husbands and wives net of the associations with the women's autonomy indices and the factors identified as associated with age at marriage in Chapter 5 (Figure 5.9). This also examines whether there is a direct association between the women's autonomy indices and the age difference, net of the association between the age difference and age at marriage, and the factors identified as having a direct association with age at marriage.

The saturated model is expressed as ghmpqrtuxyz | ab. Thus, the factors associated with age at marriage (coded g, h, m, p, q, r, t, and u) and the women's autonomy indices (coded x, y, and z) are treated the same, i.e. all in block one, while the husband's age at marriage and the wife's age at marriage are in block two. A parsimonious hierarchical interaction model is identified from the stepwise selection procedure in MIM ($p<.05$). In MIM notation, this is expressed as:

1. ghmpqrtuxyz
2. ghmpqrtuxyz/bghqruz, aghmqrtuxyz/ab

Thus, in block one, all the factors associated with age at marriage and the women's autonomy indices have direct associations with each other ($p<.05$). In block two, the wife's age at marriage (coded b) has statistically significant direct associations with five of the factors (codes g, h, q, r, and u) and the realised autonomy index (coded z). While the husband's age at marriage (coded a) has a statistically significant direct association with all of the factors except the wife's childhood place of residence (code p), and all three of the women's autonomy indices (codes x, y, z). Figure 6.7 shows this graphical chain model.

Figure 6.7 A graphical chain model of the associations between the ages at marriage of husbands and wives, women's autonomy indices and factors associated with the age at marriage: Egypt



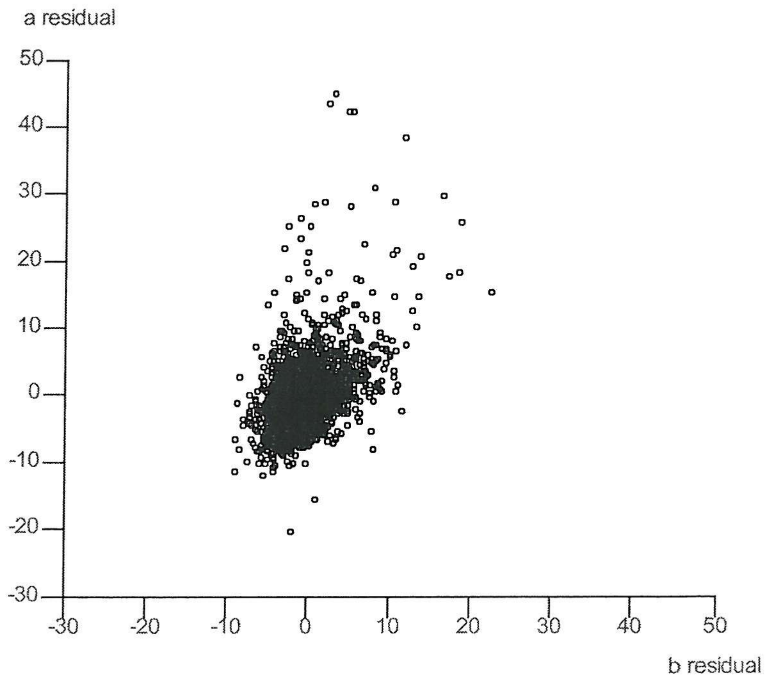
Key for Figure 6.7

a	husband's age at marriage	r	post-marital residence when first married
b	wife's age at marriage	t	wife knew & chose husband
g	wife's highest level of education	u	husband's profession
h	husband's highest level of education	x	customary autonomy index
m	consanguineous marriage	y	non-customary autonomy index
p	wife's childhood place of residence	z	realised autonomy index
q	wife worked before marriage		

While Figure 6.5 showed statistically significant direct associations between the wife's age at marriage and the customary and non-customary autonomy indices, Figure 6.7 shows that the wife's age at marriage only has a direct association with the realised autonomy index, net of the associations in and between blocks 1 and 2 in this model. In contrast, Figure 6.7 shows a direct association between the husband's age at marriage and all three of the autonomy indices net of the associations in and between blocks 1 and 2 in this model. The inclusion of the additional factors shown in Figure 6.7 seems to facilitate links between the autonomy indices and husband's age at marriage so that there is a direct association between the age difference net of the wife's age at marriage and women's autonomy, net of the associations between the autonomy indices and the factors associated with age at marriage, the associations between these indicators and the ages at marriage, and the association between the ages at marriage of husbands and wives. Thus, Figure 6.7 does not support the conclusion drawn from Figure 6.5, which showed no direct association between any of the autonomy indices and the age difference net of the wife's age at marriage, i.e. the husband's age at marriage.

As far as the hypothesis that the age difference is a by-product of factors associated with marriage timing (the hypothesis examined in Chapter 5) is concerned, the graphical chain model in Figure 6.7 shows that even after controlling for a number of factors associated with age at marriage, there remains a direct association between the ages at marriage of husbands and wives. This association is highly statistically significant (*p*=.000 according to an edge exclusion test). Figure 6.8 shows some positive albeit weak association between the partial residuals for the husband's age and wife's age for this parsimonious model. The corresponding partial correlation coefficient is .140, which compares to the Pearson correlation coefficient for this sample of .451, and a partial correlation coefficient obtained from the two separate linear regressions of .364. Thus, as concluded in Chapter 5, it seems unlikely that the association between the ages at marriage of husbands and wives is simply a by-product of factors associated with age at marriage, including hypothesised indicators of women's autonomy.

Figure 6.8 Scatterplot of the residuals from the hierarchical interaction model of the ages at marriage of husbands (a) and wives (b) and eleven factors associated with age at marriage available from the women's status module of the 1995 Egyptian DHS



6.9 Discussion and conclusions

The objective of this last chapter of analyses was to examine the hypothesis that the age difference reflects the status of women at the individual level. Chapter 3 examined this hypothesis at the aggregate level and found that, because of an association between the age difference and the average age at marriage of females, and an association between hypothesised indicators of women's status and the average age at marriage of females, no association was observed between the age difference and the women's status indicators, net of the average age at marriage of females. However, it could not be assumed that the same conclusion could be drawn at the individual level, especially since it was observed in Chapter 4 that, within countries, a greater association exists between the age difference and the husband's age at marriage than with the wife's age at marriage. Thus, this chapter examined this hypothesis at the individual level in the context of one country. Egypt was focussed upon because the 1995 Egyptian DHS survey included additional questions on women's status. As it is an Islamic society with relatively rigid norms about gender roles (Rugh, 1984, cited by Kishor, 1995), it was considered an interesting country on which to focus a culturally specific study of the association between the age difference and women's status.

As discussed in Chapter 2, women's status can, and has been interpreted in a number of different ways. The analysis in this chapter focused on one particular interpretation of women's status: women's autonomy. However, since it has been argued that there are different dimensions of autonomy (e.g. Mason, 1993; Simmons and Young, 1996), fifteen questions were identified from the women's status module, which together were hypothesised to reflect three dimensions of women's autonomy: customary autonomy, non-customary autonomy, and realised autonomy. These three dimensions were loosely quantified with three indices based on a method defined by Kishor (1995). The indices had similar distributions to those reported by Kishor (1995, p. xi), even though she analysed data from the 1988 Egyptian DHS. Some positive association was observed between the three indices but this was not perfect, which perhaps suggests that the indices are tapping different dimensions of women's autonomy as hypothesised (Balk, 1994; Kishor, 1995). This in turn supports the notion that a woman may have high status in one area of her life, but low status in another (e.g. Hagan, 1983; Schildkrout, 1983; Casterline *et al*, 1986).

Overall, women in this sample have the greatest autonomy according to the customary autonomy index, with over half of these women obtaining the highest score. These results are not really surprising given the sample selection criteria that were adopted, i.e. focussing on women married in the ten years prior to interview who had children. However, it is likely that similar results would have been observed if broader sample selection criteria had been used because, as discussed in Chapter 2, it has been argued that marriage and childbearing are central events in the lives of most Egyptian women (Nawar, 1988; Sokona and Casterline, 1988; Mahran *et al*, 1995), since *"the family unit is central in Arab societies. Societal recognition and support systems revolve around the roles of women as wives and mothers"* (Rashad, 2000, p. 94). While this emphasis on rigid gender norms and traditionalism may be interpreted as women having low status, women are likely to have considerable autonomy over reproductive issues because these issues fall within the *"socially prescribed female domain"* (Govindasamy and Malhotra, 1996, p.329). However, due to improvements in access to education for females in Egypt (UN, 1999), women may aspire to greater education for their daughters, and liberalised attitudes toward female employment may make the idea of women working outside the home more socially acceptable (Assaad, 1980). Thus, it is possible that over the next generation Egyptian women will score more highly on the non-customary and realised autonomy indices.

Of course, it is acknowledged that this method of conceptualising and quantifying women's autonomy is not definitive. Although this method has been used in a published study (Kishor, 1995) this does not equate to validation. Other studies have adopted different approaches. For example, in a cross-cultural discussion of autonomy, Jejeebhoy (1996) conceptualised autonomy as having five distinct dimensions: informative, decision-making, physical, emotional and economic and social. In contrast, Schuler *et al* (1997, p. 563) discussed women's autonomy in the context of six villages in Bangladesh in relatively more specific terms, such as women's participation in public protests and political campaigning, and the ability to make 'larger purchases'. As with Schuler *et al*'s (1997) approach, adopting Kishor's (1995) method has the advantage that autonomy has a culturally-sensitive definition, but consequently the disadvantage that the results are not generalisable to other settings. Indeed, given the subjectivity regarding the variables used to derive the indices, the coding of the responses to these variables, and the categorisation of low, medium and high levels of autonomy, it is possible to question whether re-definition may yield different results.

It is also worth noting that using Kishor's (1995) method to create the autonomy indices meant that it was necessary for women to have responded to all 15 questions used to derive the women's autonomy indices to be included in this analysis. While almost every woman who met the selection criteria for this analysis and who were eligible for the women's status module completed this module (99.3 % of the 2002 women eligible for the women's status module), 41.6% of these 2002 women did not respond to at least one of these 15 questions. This may reflect the relatively sensitive nature of the questions in the women's status module relative to the rest of the women's individual questionnaire. Thus, it is possible that the data are subject to systematic bias, i.e. women who did not respond to all 15 questions used in this analysis may have lower status relative to those women who did respond to these 15 questions. This may have limited the extent of variation in the women's autonomy indices among the 1168 women who were studied.

Another limitation is that the sample selection criteria may limit the extent to which the results can be applied to other population groups within Egypt. However, it was necessary to focus the sample in order to prevent spurious associations. For example, statistically significant variation in the age difference by level of autonomy exists for more of the questions studied when the sample consists of all women in their first marriage regardless of the number of years married ($p<.05$). However, a highly statistically significant association also exists for such questions between the number of years married and autonomy level ($p=.000$). While the problem of limited generalisability was discussed in Chapter 3 with respect to studying women's status cross-nationally, it seems that this problem is not overcome by focussing on one country. This again raises questions regarding the usefulness of women's status as a concept.

As far as the hypothesis that the age difference reflects women's status, or more specifically, women's autonomy at the individual level, this analysis observed little evidence in support of this hypothesis. When women's autonomy was considered multidimensionally in terms of the three indices of women's autonomy, little association with the age difference was observed for this sample. Variations in the average age difference were only statistically significant for the customary autonomy index and then only for women with a medium relative to a high level of autonomy and only at $p<.05$ but not $p<.01$. While it is important to consider women's status multidimensionally (e.g. Caldwell and Caldwell, 1993; Balk, 1997), it is also important to consider the individual components of such indices as hypothesised indicators of women's autonomy in their own right (e.g. Mason, 1995; Govindasamy and Malhotra, 1996). Statistically significant variation in the average age difference was observed for a couple of the 15 indicators, although the association was not in the hypothesised direction for one of the indicators. Given the inconsistency in the reporting of either a less or more autonomous response to the 15 questions, it is not surprising that there is no consistency for these variables in the statistical significance of the variation in the average age difference or in the direction of the association. It is also possible that, given the number of associations that were considered, one or indeed both of the statistically significant associations occurred by chance. This lack of association may also partly reflect relatively little variation in the women's autonomy indicators and the age difference for this sample of women. Others have investigated the association between the age difference and a number of social and economic factors and have reported that these factors only account for a small amount of the variation in the age difference (Sokona and Casterline, 1988), which had been attributed to a weak recognition of the age difference in Egyptian society beyond adhering to simple rules discouraging extreme age differences (Hallouda *et al*, 1988).

Chapter 3 identified the importance of the association between the age difference and the average age at marriage of females in considering the association between the age difference and women's status at the aggregate level. The analysis in Chapter 4 paid attention to the association between the age difference and age at marriage within countries, and observed that a greater proportion of the variation in the age difference could be accounted for by the husband's age at marriage rather than the wife's age at marriage. However, while statistically significant variations in the average age difference were only observed in this chapter for the customary autonomy index, highly statistically significant variations in the average ages at marriage of both husbands and wives were observed for all three autonomy indices.

As in Chapter 5, this chapter used graphical modelling to simultaneously consider the association between the autonomy indices and the age difference net of the age at marriage of wives, and between the autonomy indices and the age difference net of the age at marriage of husbands. As reported in the aggregate level analysis in Chapter 3, no association was observed between any of the women's autonomy indices and the husband's age at marriage, which as shown in Chapter 2, is statistically equivalent to the age difference net of the wife's age at marriage (Plewis, 1985). This suggests that within countries, or rather, at least within Egypt, the weak association between the age difference and women's status is spurious in a statistical sense due to three associations (1) between the age difference and the age at marriage of husbands; (2) between the ages at marriage of husbands and wives; (3) between the age at marriage of wives and women's status. This model also identified that net of the other associations, there remained a statistically significant association between the ages at marriage of husbands and wives, as concluded from the analysis in Chapter 5.

For completion, Chapter 6 also examined the association between the ages at marriage of husbands and wives net of the autonomy indices and the factors identified in Chapter 5 as significantly associated with age at marriage. Net of these factors a statistically significant association remained between the ages at marriage of husbands and wives, which is again consistent with Chapter 5's findings. As far as the women's status hypothesis is concerned it is interesting to note direct associations were observed from this model between the autonomy indices and the ages at marriage of both husbands and wives. Thus, these results suggest that net of other factors associated with age at marriage, the wife's autonomy maybe important for the timing of marriage for both the wife and her husband. More generally, these results suggest that in addition to an individual's own characteristics, their partner's characteristics are also important for the timing of marriage. This seems logical given the theory that individuals may seek partners with specific characteristics, (e.g. a man may desire a wife with relatively high autonomy), or the theory that individuals may seek partners with characteristics that they do not possess themselves (e.g. a woman looking for a husband of higher socio-economic status), as discussed in Chapter 1.

It is important to recall however, that the autonomy indices used in this analysis are derived from variables referring to the wife's behaviour at the time of interview rather than her behaviour before marriage. The wife's autonomy may therefore be a product of the actions of both the husband and wife. None the less, these results suggest that it is necessary to consider the characteristics of both husbands and wives in order to understand spouse choice and thus the timing of marriage. In other words, to adopt a *"two-sex perspective"* (e.g. Schoen, 1983; Xenos and Gultiano, 1992) as previously discussed.

Thus, this chapter concludes that it is unlikely that the age difference reflects women's status at the individual level in the context of Egypt, but that a woman's status maybe an important factor associated with the timing of marriage for both her *and* her husband. However, it is not possible to ascertain the direction of these associations from this analysis.

Chapter 7:
Thesis conclusions

The objective of this final chapter is to draw together the results of the analyses presented in Chapters 3 through 6 of this thesis. The thesis had two main objectives. The first objective was to examine variations in the average age difference between countries and to compare the age difference distributions for particular countries. The results of these analyses are reviewed first. The second objective of this thesis was to examine some of the hypotheses proposed in the literature for explaining these variations in the age difference. The chapter proceeds to discuss the results of these examinations, distinguishing between aggregate-level and individual-level analyses, and relating them to the results of other studies. The chapter ends with some recommendations for further investigation.

Chapter 3 began with a descriptive analysis of cross-national variations in the average age difference using aggregate level data from the United Nations' Women's Indicators and Statistics ('WISTAT') database for 93 countries. In this analysis, the average age difference was calculated as the difference between the average ages at first marriage of males and females, which were estimated using the singulate mean age at marriage ('SMAM'). Cross-national variations in the average age difference were also analysed using individual level data for a sample of 30 less developed countries from Demographic and Health Surveys ('DHS'). In the analysis presented in Chapter 4, the average age difference was calculated from individual age differences as the difference between the ages at marriage of each husband and his wife. The analysis of DHS data took a lead from a study by Casterline *et al* (1986) of World Fertility Survey ('WFS') data, so the same sample selection criteria were used in order to permit comparisons. This meant that, unlike the WISTAT estimates, the DHS estimates referred to a cohort of women who were in their first marriage or cohabitation, which began in the ten years before interview, and thus husbands may have been previously in union.

Regardless of the cultural setting, it was observed in Chapters 3 and 4 that the age difference is positive at the aggregate level. While data were only available for a sample of countries at both levels of analysis, this result is consistent with the findings of other studies as discussed in Chapter 1 (e.g. Dixon, 1971; Goldman *et al*, 1984; Bozon, 1991; Ní Bhrolcháin, 1992, 1997). Regional patterns in the average age difference were observed. Again, while the regional samples used in this thesis were incomplete, patterns observed at the aggregate- and individual-levels were similar, thus regardless of how the average age difference was measured, i.e. as the difference in SMAMs or as the age difference between individual couples. Relative to countries in other regions, age differences in sub-Saharan Africa tend to be statistically significantly larger, on average. This is especially so in the West African countries studied where average age differences are often greater than 8 years. Chapter 4 observed that, despite the tendency for age differences to be larger on average for polygamous couples relative to monogamous couples, and a relatively high prevalence of polygamy in the sub-Saharan African countries, age differences for monogamous couples in these countries still tend to be significantly larger than for monogamous couples in other regions. These results are consistent with the findings of other studies (e.g. Goldman and Pebley, 1989; Timaeus and Reynar, 1998).

In both Chapters 3 and 4, Bangladesh was identified as an interesting country in terms of the average age difference as it was the only one of the countries studied outside of sub-Saharan Africa with an average age difference of the magnitude observed in West Africa. This was estimated from the difference in SMAMs as 7.5 years while the median age difference was calculated as 7.7 years. There was much variation in the average age difference between the South Asian countries studied in both data sources. In contrast, average age differences among the Central and South America countries studied were relatively small, typically between two and four years, and thus similar in magnitude to the average age differences observed for some of the more developed countries. Average age differences for the Central and South American countries were estimated to be slightly larger from the DHS data relative to the estimates from the WISTAT data. This reflects how the DHS samples included women who were either in their first marriage or cohabitation, and that the prevalence of cohabitation has been reported as relatively high in Central and South American societies (e.g. see Appendix 4A, also Greene and Rao, 1995; Parrado and Tienda, 1997). However, evidence from French data suggest that age differences tend to be smaller, on average, for cohabiting couples relative to married couples (Bozon, 1991). Either way, it has been suggested that the variations in the age difference by union status are due to variations in the age at the start of the union for those cohabiting relative to those marrying (Ní Bhrolcháin, 1997). Chapters 3 and 4 also examined this important, yet often overlooked association between the age difference and age at marriage. Consistent with the findings of other studies, this thesis observed a greater association between the average age difference and the average age at which females marry as opposed to the average age at which males marry (e.g. Bogue, 1969; Nawar, 1988; Ní Bhrolcháin, 1992, Bozon, 1991).

Given the greater variation in the average age difference between and within less developed regions in comparison to the relatively more developed regions observed in Chapter 3 (as well as previously reported by for example, the United Nations, 1990) Chapter 4 used individual level data from the DHS surveys to examine the distributions of age differences within 30 less developed countries (the total number of countries with data available at the time of data collation). For all 30 countries, Chapter 4 reported evidence of much within-sample variation in the age difference as measured by the inter-quartile range ('IQR'). In addition, variation exists between countries in terms of the magnitude of the IQR, ranging from a few years to more than ten years. In general, IQRs tend to be larger for countries with larger average age differences, i.e. the sub-Saharan African countries and Bangladesh in this sample. It is acknowledged that studying just 30 countries limits the extent to which the results can be generalised, however, these results are consistent with those reported by Casterline *et al* (1986).

The association between the age difference and the age at marriage of in turn, husbands and wives was examined for each of the 30 countries. Contrary to the cross-national observation, a greater association was observed between the age difference and the husband's age at marriage within-countries. This result reflects greater variation in the age at marriage of men relative to women within countries. These results are consistent with findings reported by Casterline *et al* (1986), and as Chapter 4 showed, are not due to the selection criteria applying only to women.

The second objective of this thesis was to examine three hypotheses proposed and sometimes merely assumed in the literature for explaining the variations in the average age difference and its distributions. These hypotheses were that the variations:

1. reflect the age-sex structure of the eligible men and women in the marriage market;
2. are a by-product of factors associated with age at marriage;
3. reflect the status of women, which is associated with the idea that there are preferences regarding the age difference..

Chapter 1 cited numerous studies at both the aggregate and individual levels that had considered the hypothesis that the age difference reflects women's status. Despite the popularity of this hypothesis, few of these studies had actually examined the association between the age difference and indicators of women's status, let alone this hypothesised association net of the association that exists between female age at marriage and the age difference. As discussed in Chapter 2, this may partly be due to the continued debate regarding how women's status should be conceptualised and measured such that there is no definitive measure of women's status. In recognition of this ambiguity, the aggregate level analysis in Chapter 3 used several contrasting approaches to define women's status. Similarly, the individual level analysis in Chapter 6 used both unidimensional and multidimensional indicators as defined by Kishor (1995) in an attempt to capture one specific interpretation of women's status, women's autonomy. Evidence of the hypothesised negative association was observed between some of the hypothesised indicators of women's status and the age difference at both levels of analyses. However, this association was generally weak and disappeared when the linear effect of female age at marriage was removed. These analyses suggest that nothing is added, at least in a statistical sense, by using the age difference relative to the female age at marriage to reflect women's status, at either the aggregate or individual levels. However, at both levels of analysis, it is important to acknowledge the interplay between the age difference, age at marriage and women's status, and that it is not possible to tease apart these associations in any causal sense from these analyses. Furthermore, from the individual-level analysis it is also important to consider the interplay between these variables and other factors associated with age at marriage, which may be important for the marriage timing for both husbands and wives.

Given this interplay, Chapter 5 examined the hypothesis that variations in the age difference within countries are a by-product of factors associated with the timing of marriage since this hypothesis assumed that the pattern of association between the husband's age and the wife's age is responsible for the systematic component of the distribution of age differences, i.e. the difference between the observed distribution and the distribution expected given the 'random' matching of spouses with respect to their ages at marriage. As Chapter 1 discussed, there are numerous factors associated with age at marriage, and more generally, spouse choice. Unlike the study by Casterline *et al* (1986), the examination of this hypothesis involved considering both marginal and conditional associations between as many as eleven such factors (depending on availability) and the ages at marriage of husbands and wives, simultaneously. Additionally, some data were available from the Egyptian DHS survey that referred specifically to the time of marriage rather than to the time of interview as was the case for the other DHS datasets. However, for all countries studied, a highly statistically significant association was observed between the ages at marriage of husbands and wives (p=.000) net of the associations with and among these various factors. Thus, despite this thesis's methodological improvements, the same conclusion is reached as reported by Casterline *et al* (1986). The association between the ages at marriage of husbands and wives is unlikely to be simply a by-product of factors associated with the timing of marriage.

In addition to presenting a descriptive analysis of the extent of cross-national and within country variation in the age difference and age at marriage, Chapter 4 also examined the marriage market hypothesis. This thesis could have used the sex ratio of men in an eligible age range to women in an eligible age range in an attempt to quantify the marriage market at the aggregate level. However, while some studies have used this method (e.g. Akers, 1967; Dixon, 1971), it has often been criticised for being too crude as Chapter 1 discussed. Instead, joint distribution theory was used with the individual-level data to examine the marriage market hypothesis within countries, an approach used by Casterline *et al* (1986). Although, as acknowledged in Chapter 4, the DHS data refer to individual women and their husbands, so these samples may not be representative of the whole marriage market. Bearing this limitation in mind, Chapter 4 concluded that the observed age differences were, on average, of the expected magnitude given the distributions of ages at marriage for the husbands and wives in each sample. This is an important result because it contradicts Casterline *et al*'s (1986) finding that the expected average age differences were homogeneous, at around five or six years for all 28 countries that they studied. This is the key difference between the two studies since Casterline *et al* (1986) rejected the marriage market hypothesis on this evidence. However, in agreement with the results reported by Casterline *et al* (1986) is the evidence that relatively more extreme age differences are not as common as expected since the observed distributions of age differences were found to be more concentrated around the average than expected for all samples studied. Furthermore, for almost all of the countries studied, negative age differences were observed much less frequently than expected relative to 'larger' age differences, such as those in excess of ten or 15 years. Casterline *et al* (1986) interpreted these results as evidence of a tendency for individuals to avoid extreme age differences, and that 'larger' age differences are more acceptable than negative age differences in most societies.

This thesis does not assume causality, thus in this context, it does not agree with Casterline *et al*'s (1986) conclusions regarding the consequential acceptance of the preferences hypothesis, especially since this hypothesis can only be examined with explicit data on preferences (Ní Bhrolcháin, 2000). Since this thesis accepts that the empirical findings clearly reflect more than straightforward consequences of the age-sex structure, one recommendation for further study would be to obtain data on age preferences, such as from dating agencies as Ní Bhrolcháin (1997) reported in the context of Britain. Data on age preferences would be particularly interesting in the context of more traditional settings given the continuing change to union formation (e.g. Lindenbaum, 1981; Isuigo-Abanihe, 1994). If evidence was found in support of the preferences hypothesis then a further recommendation would be to examine the hypothesis that the preferences in these settings were flexible and that the marriage market operates in *"fluid way"* as Ní Bhrolcháin (2000, p.1) proposed from an analysis of British cohort data. However, the availability of data for both recommendations, especially in terms of comparable data for different settings, is unfortunately questionable at the time of writing. The DHS survey programme already asks a number of nuptiality questions so the inclusion of questions on preferred spouse age may be possible. However, since the target population of the DHS is largely, and sometimes, exclusively currently married women, it may be difficult to obtain reliable data on preferences since the questions may be answered hypothetically as many women would already be in union, reflecting the continuing importance of marriage in such societies since as Fricke *et al* (1986) argued: *"few institutions are as pivotal as marriage in their implications for a broad range of social processes."*

Appendix 2A

Countries in the WISTAT3 sample by region as defined by the United Nations Statistical Division

West Africa (n=6)	East, Central, & South Africa (n=12)	North Africa & West Asia (n=12)	South Asia (n=6)	East Asia (n=11)	Central America (n=9)	South America (n=10)	developed countries (n=27)
Benin	Central African Republic	Algeria	Bangladesh	China	Costa Rica	Argentina	Australia
Guinea Bissau	Comoros	Bahrain	India	Fiji	Cuba	Brazil	Austria
Mauritania	Congo	Egypt	Myanmar	Hong Kong*	Guatemala	Chile	Belgium
Niger	Ethiopia	Iraq	Nepal	Indonesia	Haiti	Colombia	Bulgaria
Senegal	Mauritius	Israel	Pakistan	Korea (Republic of)	Jamaica	Ecuador	Canada
Sierra Leone	Mozambique	Jordan	Sri Lanka	Malaysia	Mexico	Guyana	Czechoslovakia (former)*
	Reunion	Kuwait		Papua New Guinea	Panama	Paraguay	Denmark
	Somalia	Morocco		Philippines	Puerto Rico	Peru	Finland
	South Africa	Oman		Singapore	Trinidad & Tobago	Uruguay	France
	Swaziland	Saudi Arabia		Thailand		Venezuela	Greece
	Zaire	Syrian Arab Republic		Viet Nam			Hungary
	Zambia	United Arab Emirates					Ireland
							Italy
							Japan
							Netherlands
							New Zealand
							Norway
							Poland
							Portugal
							Spain
							Sweden
							Switzerland*
							Turkey
							United Kingdom
							United States
							USSR (former)*
							Yugoslavia (Socialist Federal Republic)*

* denotes country is not a member of the UN as at 13 March 2001

Appendix 2A (continued)

Countries excluded from the WISTAT3 sample by region as defined by the United Nations Statistical Division (n=102)

West Africa (n=9)	East, Central, & South Africa (n=21)	North Africa & West Asia (n=17)	South Asia (n=3)	East Asia & the Pacific (n=13)	Central America (n=13)	South America (n=2)	developed countries (n=24)
Burkina Faso	Angola	Afghanistan	Bhutan	Kiribati	Antigua & Barbuda	Bolivia	Albania
Cameroon	Botswana	Armenia	Cambodia	Korea (Democratic People's Republic)	Bahamas	Suriname	Andorra **
Cape Verde	Burundi	Azerbaijan	Maldives	Lao (People's Democratic Republic)	Barbados		Belarus
Côte d'Ivoire	Chad	Brunei Darussalam		Marshall Islands	Belize		Bosnia & Herzegovina
Ghana	Djibouti	Cyprus		Micronesia (Federated States of)	Dominica		Croatia
Guinea	Equatorial Guinea	Georgia		Mongolia	Dominican Republic		Czech Republic
Mali	Eritrea	Iran (Islamic Republic)		Nauru**	El Salvador		Estonia
Nigeria	Gabon	Kazakhstan		Palau	Grenada		Germany
Togo	Gambia	Krygyzstan		Samoa	Honduras		Iceland
	Kenya	Lebanon		Solomon Islands	Nicaragua		Latvia
	Lesotho	Liberia		Tonga	Saint Kitts & Nevis		Liechtenstein
	Madagascar	Libyan Arab Jamahiriya		Tuvalu**	Saint Lucia		Lithuania
	Malawi	Qatar		Vanuatu	Saint Vincent & the Grenadines		Luxembourg
	Namibia	Tunisia					Macedonia (the former Yugoslavia Republic)
	Rwanda	Tajikistan					Malta
	Sao Tome & Principe	Turkmenistan					Monaco
	Seychelles	Uzbekistan					Moldova (Republic)
	Sudan						Romania
	Uganda						Russian Federation
	United Republic of Tanzania						San Marino
	Zimbabwe						Slovakia
							Slovenia
							Ukraine
							Yugoslavia

** Data not presented for this country/area in WISTAT3

Appendix 2A (continued)

Countries excluded from the WISTAT3 sample used in Chapter 3 that are in WISTAT3 but which are not UN members as at 13 March 2001 (n=20)

American Samoa
Cook Islands
East Timor
French Guyana
French Polynesia
German Democratic Republic
German Federal Republic
Guadeloupe
Guam
Macau
Martinique
Netherlands Antilles
New Caledonia
Northern Mariana Islands
Pacific Islands
US Virgin Islands
Western Sahara
Yemen (former Democratic Republic)
Yemen (former)
Yugoslavia (former Republic)

Source: United Nations (1997) Composition of macro geographical (continental) regions and component geographical regions. http://www.un.org/Depts/unsd/methods/m49regin.htm. New York: United Nations Statistical Division

Appendix 2B

Data availability for the hypothesised indicators of women's status from the WISTAT3 used for the factor analysis and supplementary data sources

Indicator	% of the sample of 93 countries with data available from WISTAT3	Definition
Adult illiteracy [1]	81.7	Proportion of the population aged 15 and over who are illiterate or semi-illiterate
Secondary school enrolment [2]	94.6	Population aged 12-17 enrolled in school divided by the corresponding population.
Female economic activity	100.0	Percent of females aged 20-24 who are economically active defined by the International Labour Office as *'persons working for pay or profit any time during a specified reference period or seeking such work'*. (United Nations, 1994, p.16)
Female household headship [3]	75.3	Per cent of all household heads who are female
Total fertility rate	100.0	Average number of children borne to a woman
Contraceptive use [4]	87.1	Per cent of women using any contraceptive method
Maternal mortality	100.0	Maternal mortality rate per 100,000 live births
Extent of safe childbirth [5]	95.7	Per cent of births attended by a trained birth attendant
Life expectancy at birth	100.0	Number of years newborn males or females would be expected to live according to the age-specific mortality rates prevailing at the time of their birth
Representation of women in parliamentary assembly [6]	92.5	Proportion of the lower chamber or unicameral assembly seats occupied by women
Representation of women in ministerial posts [7]	93.5	Percentage of ministerial posts held by women

Notes for Appendix 2B:

1. Data missing from WISTAT3 for Oman are supplemented by data available at the United Nations Statistical Division website for Oman. The other 16 countries for which data are missing are all developed and so it is assumed that adult illiteracy is negligible in these populations relative to less developed populations. These countries are: Australia, Austria, Belgium, Canada, former Czechoslovakia, Denmark, Finland, France, Ireland, Japan, Netherlands, New Zealand, Norway, Sweden, Switzerland, and the United Kingdom.
2. Data missing from WISTAT3 are supplemented by 1985 data available at the United Nations Statistical Division website for Congo, Jordan, Puerto Rico, Reunion, and Switzerland.
3. Data missing from WISTAT3 are supplemented for Senegal by 1985 data available at the United Nations Statistical Division website. Regional medians are used as best estimates for the other 22 countries.
4. Data missing from WISTAT3 are supplemented for C.A.R., Chile, Comoros, United Arab Emirates, and Zaire by 1985 data available at the United Nations Statistical Division website. For the other countries: Congo, Greece, Ireland, Israel, Saudi Arabia, Russia, and Uruguay regional medians are used as best estimates. However, it is acknowledged that the extent to which within-region homogeneity can be assumed is questionable.
5. Data missing from WISTAT3 are supplemented for Austria and Ireland by data from the World Health Organisation (1991) *Coverage of maternity care: A listing of information available.* For the Congo and Zaire, regional medians are used as best estimates.
6. Data missing from WISTAT3 are supplemented for Ethiopia and Kazakhstan by data from Inter-Parliamentary Union (1997) *Men and women in politics: Democracy still in the making.* For Bahrain, Oman, Puerto Rico, Reunion and Saudi Arabia, regional medians are used as best estimates.
7. Data missing from WISTAT3 are supplemented for Kazakhstan, Russia, and Yugoslavia by data available at the United Nations Statistical Division website. For the other countries: Czechoslovakia, Puerto Rico, and Reunion regional medians are used as best estimates.

Appendix 2C

Percent of women in each DHS sample who were cohabiting at interview

Country by region [1]	Target population [2]	Sample used for this analysis:	
		% cohabiting	Weighted n
Central & South America			
Brazil	AW (de jure) 15-49	25.4	2716
Colombia	AW (de jure) 15-49	51.2	2500
Dominican Republic	AW 15-49	57.2	1765
Guatemala	AW 15-49	40.0	2954
Haiti	AW 15-49	51.1	548
Peru	AW 15-49	53.8	7023
West Asia & North Africa			
Egypt	EMW 15-49	0.0	5344
Kazakhstan	AW 15-49	2.8	933
Uzbekistan	AW 15-49	0.6	1383
South Asia			
Bangladesh	EMW 10-49	0.0	2551
Nepal	EMW 15-49	0.0	3008
Pakistan	EMW 15-49	0.0	2474
Philippines	AW 15-49	10.5	3650
Central, East & South Africa			
Central African Republic	AW 15-49	84.7	1381
Comoros	AW 15-49	2.2	502
Kenya	AW 15-49	5.0	1270
Malawi	AW 15-49	4.4	997
Mozambique	AW 15-49	78.8	2212
Rwanda	AW 15-49	47.3	1400
Tanzania	AW 15-49	9.4	2211
Uganda	AW 15-49	10.4	1860
Zambia	AW 15-49	1.1	2126
Zimbabwe	AW 15-49	0.0	923
West Africa			
Benin	AW 15-49	11.9	1615
Burkina Faso	AW 15-49	0.0	1757
Cameroon	AW 15-49	10.6	683
Côte d'Ivoire	AW 15-49	0.0	1203
Ghana	AW 15-49	12.1	660
Mali	AW 15-49	2.5	2661
Niger	AW 15-49	0.0	1150

Notes for Appendix 2C
1. Countries are shown in alphabetical order
2. AW = all women; EMW = ever married women

Appendix 2D

Age ratios for each age, by country according to the more reliable data source [1]: Husbands

Country	Source [2]	Husband's age [3]																			Total misreported	Mean absolute deviation
		20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38		
Philippines	WQ				101	110	94	104	96	101	110	92	102	104	100						**0**	**5.5**
Brazil	WQ					101	99	97	100	108	96	105	92	**116**	89	100					**1**	**4.9**
Colombia	WQ					101	109	91	104	110	86	**126**	**77**	100	111	99					**2**	**9.8**
Mozambique	WQ				105	94	110	98	100	102	88	**141**	**59**	113	94						**2**	**12.6**
Kenya	MS						113	105	**76**	**121**	98	109	89	109	99	107	**76**				**3**	**11.4**
Peru	WQ				103	102	101	101	95	115	**82**	**131**	**76**	111	98						**3**	**12.5**
Côte d'Ivoire	HHS								**134**	**70**	94	**124**	99	95	85	110	106	100	86		**3**	**13.3**
Uzbekistan	WQ			99	104	96	**117**	88	110	105	**82**	**124**	88	114							**3**	**10.9**
Kazakhstan	WQ				**116**	90	111	96	104	90	**118**	98	87	**118**	92						**3**	**10.5**
Dominican Republic	WQ					103	94	109	97	105	**83**	**138**	**70**	111	112	**78**					**4**	**14.3**
C.A.R.	HHS				100	104	105	98	93	102	99	**124**	**72**	**123**	**75**						**4**	**11.0**
Zimbabwe	HHS					**144**	102	**80**	113	90	103	112	97	106	**70**	**143**					**4**	**14.4**
Comoros	HHS									111	**84**	**126**	**83**	**124**	87	77	**132**	94	85	95	**5**	**17.2**
Guatemala	WQ		85	113	98	100	112	**83**	**117**	96	**82**	**147**	**59**								**5**	**16.8**

Appendix 2D (continued)

Age ratios for each age, by country according to the more reliable data source [1]: Husbands (continued)

Country	Source [2]	Husband's age [3]																			Total misreported	Mean absolute deviation
		20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38		
Cameroon	HHS							75	102	105	101	107	88	132	64	123	84	123			6	16.6
Burkina Faso	HHS								105	104	79	141	73	114	107	66	143	81	114		6	20.8
Rwanda	WQ					35	167	68	119	102	95	100	115	85	130	43					6	27.9
Zambia	HHS				95	117	86	107	90	133	76	110	82	133	75						6	17.7
Ghana	HHS							97	99	115	66	162	56	142	76	73	146	91			7	27.8
Haiti	HHS						100	92	132	73	97	135	53	155	80	88	142				7	25.7
Nepal	WQ	120	93	118	79	101	122	89	97	118	62	155									7	19.4
Niger	HHS						152	81	107	108	71	160	61	107	113	67	150				7	28.8
Egypt	WQ							83	103	112	74	142	76	116	94	79	147	73			8	22.4
Mali	HHS								94	128	58	170	59	124	91	63	172	62	93		8	33.9
Malawi	HHS				71	118	91	92	77	163	61	129	71	143	63						9	29.6
Uganda	HHS			100	113	81	136	83	87	130	66	156	61	140							9	27.0
Benin	HHS						133	90	79	152	48	172	55	139	88	70	140				9	36.9
Tanzania	WQ						123	96	88	130	61	167	55	138	80	78	146				9	31.3
Pakistan	MS					65	151	90	77	140	53	187	38	153	84	75					10	40.8
Bangladesh	HHS						156	83	92	128	41	216	30	175	68	35	223				10	59.0
Number of times age is:																						
Under-reported		0	0	0	1	3	0	7	4	1	17	0	19	0	7	2	2	3	0	0	71	
Over-reported		1	0	1	2	2	8	0	3	8	1	22	0	15	1	10	10	1	0	0	90	

Appendix 2D (continued)

Age ratios for each age, by country according to the more reliable data source [1]: Wives

Country	Source [2]	Wife's age [3] 15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	Total misreported	Mean absolute deviation
Brazil	WQ							101	100	101	97	105	97	108	99	107	94	98	**0**	**3.6**
Dominican Republic	WQ					96	97	103	99	101	106	97	103	98	102	94			**0**	**3.0**
Guatemala	WQ				115	96	105	98	98	109	96	111	86	99	109				**0**	**6.9**
Peru	WQ						101	101	102	98	105	98	104	97	104	97	99		**0**	**2.4**
Philippines	WQ						**77**	92	99	99	104	97	100	104	106	95	99		**1**	**3.7**
Uzbekistan	WQ					95	111	102	95	107	95	110	100	97	109	**72**			**1**	**7.7**
Bangladesh	WQ	96	**118**	90	109	93	105	104	105	96	93	105							**1**	**7.1**
Colombia	WQ						**117**	85	106	96	110	101	97	102	100	96	111		**1**	**6.6**
Nepal	WQ			93	111	93	113	87	**117**	89	113	96	100	100					**1**	**8.8**
Tanzania	WQ				104	98	109	101	105	89	107	109	94	94	**120**				**1**	**7.2**
Kazakhstan	WQ							101	95	113	93	102	95	103	103	100	**119**	**78**	**2**	**7.2**
Benin	WQ				**133**	85	104	99	105	111	**78**	**122**	99	96	113				**3**	**11.9**
Burkina Faso	WQ			111	103	95	115	88	**119**	86	94	**123**	86	**121**					**3**	**13.0**
Zambia	WQ			87	107	103	106	95	106	95	**120**	**78**	**117**	89					**3**	**10.3**
Malawi	WQ			105	109	90	**121**	**82**	108	95	**116**	90	108	**74**					**4**	**12.4**
Mozambique	WQ			87	**124**	**80**	**128**	**84**	109	99	95	110	94	96					**4**	**12.4**
Rwanda	WQ					**68**	108	92	108	110	97	86	**116**	95	**124**	**47**			**4**	**16.5**
Uganda	WQ			89	113	94	**120**	**78**	**120**	94	95	113	100	**81**					**4**	**12.2**

Appendix 2D (continued)

Age ratios for each age, by country according to the more reliable data source [1]: Wives (continued)

Country	Source [2]	Wife's age [3]																	Total misreported	Mean absolute deviation
		15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31		
Comoros	WQ					97	106	97	122	76	107	101	113	82	131	67			5	14.7
C.A.R.	WQ			92	128	80	128	75	117	93	106	106	98	85					5	14.6
Ghana	WQ						169	60	112	90	95	131	85	94	118	80	114		5	21.9
Zimbabwe	WQ					118	76	107	122	74	115	96	118	88	107	104			5	13.1
Kenya	WQ				91	78	144	84	88	121	97	98	92	100	133				5	18.1
Egypt	WQ					71	146	70	119	99	82	137	77	104	105	86			7	20.2
Mali	WQ		94	102	120	72	144	69	125	93	88	141	57						7	23.5
Niger	WQ	103	95	91	133	56	163	61	133	90	83	138							7	26.8
Côte d'Ivoire	WQ			81	136	80	126	77	128	74	118	105	100	102					8	18.3
Cameroon	WQ			58	152	74	128	81	116	89	101	120	80	95					8	21.7
Haiti	WQ						107	75	131	79	96	143	57	117	124	62	104		8	23.3
Pakistan	WQ				168	41	197	42	144	88	69	159	79	76	141				10	46.8
Number of times age is																				
Under-reported:		0	0	2	0	10	2	13	0	4	4	1	5	4	0	5	0	1	56	
Over-reported:		0	1	0	8	1	13	0	13	1	3	9	3	2	7	0	1	0	62	

Note for Appendix 2D:

1. Owing to space limitations age ratios are only shown for the more reliable data source as identified in Appendix 2E.
2. 'WQ' refers to 'Women's Questionnaire', 'HHS' refers to 'Household Schedule', and 'MS' refers to 'Men's Survey' (only applies to data for husband's age for Kenya and Pakistan).
3. Underlined age ratios correspond to the median age for that sample.

Appendix 2E

Mean absolute deviation (from 100) for each of the two data sources available from the DHS

Wife's age

Country by region	Mean absolute deviation for each source: Women's questionnaire	Household schedule
Central & South America		
Brazil	3.6	7.0
Colombia	6.6	9.4
Dominican Republic	3.0	5.7
Guatemala	6.9	13.4
Haiti	23.3	25.9
Peru	2.4	6.8
West Asia & North Africa		
Egypt	20.2	22.3
Kazakhstan	7.2	9.8
Uzbekistan	7.7	7.8
South Asia		
Bangladesh	7.1	7.4
Nepal	8.8	9.5
Pakistan	46.8	50.6
Philippines	3.7	3.4
Central, East & South Africa		
Central African Republic	14.6	16.0
Comoros	14.7	17.5
Kenya	18.1	15.7
Malawi	12.4	6.9
Mozambique	12.4	14.8
Rwanda	16.5	18.1
Tanzania	7.2	10.4
Uganda	12.2	20.5
Zambia	10.3	9.2
Zimbabwe	13.1	13.9
West Africa		
Benin	11.9	13.9
Burkina Faso	13.0	15.5
Cameroon	21.7	26.1
Côte d'Ivoire	18.3	20.5
Ghana	21.9	22.4
Mali	23.5	24.4
Niger	26.8	44.0

Appendix 2E (continued)

Husband's age

Country by region	Mean absolute deviation for each source	
	Women's questionnaire	Household schedule
Central & South America		
Guatemala	16.8	-
Peru	12.5	17.0
Brazil	4.9	5.4
Colombia	9.8	-
Dominican Republic	14.3	23.1
Haiti	-	25.7
West Asia & North Africa		
Kazakhstan	10.5	-
Uzbekistan	10.9	-
Egypt	22.4 (1)	22.1
South Asia		
Philippines	5.5	-
Nepal	19.4	-
Pakistan [2]	76.4	40.8
Bangladesh	64.3	59.0
Central, East & South Africa		
Rwanda	27.9	-
Uganda	-	27.0
Malawi	-	29.6
Mozambique	12.6	19.7
Central African Republic	18.0	11.0
Zambia	19.2	17.7
Zimbabwe	-	14.4
Kenya [2]	-	11.4
Tanzania	31.3 (1)	25.6
Comoros	24.6	17.2
West Africa		
Ghana	-	27.8
Benin	44.0	36.9
Niger	29.2	28.8
Côte d'Ivoire	-	13.3
Cameroon	18.6	16.6
Burkina Faso	-	20.8
Mali	43.7	33.9

Notes for Appendix 2E

1. No significant difference was found (at $p<.05$) between the mean absolute deviation calculated using data from the women's questionnaire data and the mean absolute deviation calculated from using data from the household schedule, thus the women's questionnaire is the preferred data source for husband's age as this results in a larger sample size than if the household schedule is used.
2. Data come from the men's survey rather than the household survey for Kenya and Pakistan.

Appendix 2F

Quality of age data according to the mean absolute deviation for preferred data source [1]

Country and extent of misreporting	Preferred data source [1, 2] and corresponding mean absolute deviation			
	Wives		Husbands	
'No misreporting' **(both mean absolute deviations <15 [3])**				
Peru	2.4	WQ	12.5	WQ
Dominican Republic	3.0	WQ	14.3	WQ
Brazil	3.6	WQ	4.9	WQ
Philippines	3.7	WQ	5.5	WQ
Colombia	6.6	WQ	9.8	WQ
Kazakhstan	7.2	WQ	10.5	WQ
Uzbekistan	7.7	WQ	10.9	WQ
Mozambique	12.4	WQ	12.6	WQ
Zimbabwe	13.1	WQ	14.4	HHS
Central African Republic	14.6	WQ	11.0	HHS
'Moderate misreporting' **(one or both mean absolute deviations ≥15 & <25 [3])**				
Guatemala	6.9	WQ	16.8	WQ
Nepal	8.8	WQ	19.4	WQ
Zambia	10.3	WQ	17.7	HHS
Burkina Faso	13.0	WQ	20.8	HHS
Comoros	14.7	WQ	17.2	HHS
Kenya [4]	18.1	WQ	11.4	MS
Côte d'Ivoire	18.3	WQ	13.3	HHS
Egypt	20.2	WQ	22.4	WQ
Cameroon	21.7	WQ	16.6	HHS
'Severe misreporting' **(one or both mean absolute deviations ≥25 [3])**				
Bangladesh	7.1	WQ	59.0	HHS
Tanzania	7.2	WQ	31.3	WQ
Benin	11.9	WQ	36.9	HHS
Uganda	12.2	WQ	27.0	HHS
Malawi	12.4	WQ	29.6	HHS
Rwanda	16.5	WQ	27.9	WQ
Ghana	21.9	WQ	27.8	HHS
Haiti	23.3	WQ	25.7	HHS
Mali	23.5	WQ	33.9	HHS
Niger	26.8	WQ	28.8	HHS
Pakistan [4]	46.8	WQ	40.8	HHS

Notes for Appendix 2F

1. See Appendix 2E.
2. 'WQ' refers to 'Women's Questionnaire' and 'HHS' refers to 'Household Schedule'.
3. The thresholds for defining 'moderate' and 'severe' misreporting are arbitrary.
4. Husband's age data from the men's survey rather than the household schedule for Kenya and Pakistan.

Appendix 2G

Data availability for husband's age for women who meet the sample selection criteria, by sample

Country by region	Unweighted no. of eligible women[1]	Husband's age obtainable from both WIQ & HHS[2]	More reliable data source[3]	No. of women with husband's age unavailable from more reliable data source[3]	No. of women with husband's age unavailable from more reliable data source[3] but with husband's age available from the alternative data source	Of remaining women with husband's age unavailable, no. of women for whom husband's age is available by linking to the HHS via the wife's relationship to the household head[4]	Percent of all eligible women for whom the husband's age is available
Central & South America							
Brazil	2720	Both	WIQ	14	11	0	99.9
Colombia	2523	WIQ only	WIQ	4	0	4	100.0
Dominican Republic	1776	Both (HHS for some only)	WIQ	7	5	0	99.8
Guatemala	3041	WIQ only	WIQ	19	0	4	99.5
Haiti	1040	HHS only (for some)	HHS	864	0	366	51.9
Peru	7314	Both (HHS for some only)	WIQ	13	0	6	99.9
West Asia & North Africa							
Egypt	5352	Both	WIQ	3	2	0	100.0
Kazakhstan	948	WIQ only	WIQ	0	n.a.	n.a.	100.0
Uzbekistan	1363	WIQ only	WIQ	3	0	1	99.9

Appendix 2G (continued)

Data availability for husband's age for women who meet the sample selection criteria, by sample

Country by region	Unweighted no. of eligible women[1]	Husband's age obtainable from both WIQ & HHS[2]	More reliable data source[3]	No. of women with husband's age unavailable from more reliable data source[3]	No. of women with husband's age unavailable from more reliable data source[3] but with husband's age available from the alternative data source	Of remaining women with husband's age unavailable, no. of women for whom husband's age is available by linking to the HHS via the wife's relationship to the household head[4]	Percent of all eligible women for whom the husband's age is available
South Asia							
Bangladesh	3231	Both	HHS	656	0	6	79.9
Philippines	3695	WIQ only	WIQ	8	0	7	100.0
Pakistan	2470	Both (Men's survey)	Men's survey	1960	1930	0	98.8
Nepal	2999	WIQ only	WIQ	3	0	0	99.9
East, Central & South Africa							
Central African Rep.	1378	Both	HHS	267	262	0	99.6
Comoros	513	Both	HHS	104	93	0	97.9
Kenya	1871	Men's survey only	Men's survey	1334	0	692	65.7
Malawi	1286	HHS only	HHS	929	0	708	82.8
Mozambique	2333	Both (HHS for some only)	WIQ	195	28	71	95.8
Rwanda	1483	WIQ only	WIQ	1245	0	1158	94.1
Tanzania	2164	Both	WIQ	25	9	0	99.3
Uganda	2184	HHS only	HHS	403	0	21	82.5
Zambia	2118	Both	HHS	265	260	0	99.8
Zimbabwe	1554	HHS only	HHS	1249	0	572	56.4

Appendix 2G (continued)

Data availability for husband's age for women who meet the sample selection criteria, by sample

Country by region	Unweighted no. of eligible women[1]	Husband's age obtainable from both WIQ & HHS[2]	More reliable data source[3]	No. of women with husband's age unavailable from more reliable data source[3]	No. of women with husband's age unavailable from more reliable data source[3] but with husband's age available from the alternative data source	Of remaining women with husband's age unavailable, no. of women for whom husband's age is available by linking to the HHS via the wife's relationship to the household head[4]	Percent of all eligible women for whom the husband's age is available
West Africa							
Burkina Faso	2040	HHS only	HHS	398	0	0	80.5
Benin	1633	Both	HHS	341	335	0	99.6
Côte d'Ivoire	2076	HHS only	HHS	1628	0	767	58.5
Cameroon	998	Both (WIQ for some only)	HHS	655	0	338	68.2
Ghana	1088	HHS only	HHS	428	0	1	60.7
Mali	2668	Both	WIQ	415	376	0	98.5
Niger	1536	Both	HHS	1083	2	679	74.3

Notes for Appendix 2G

1. Women are eligible for the sample if they were in their first marriage or cohabitation at interview, which began in the 10 years prior to interview
2. 'WIQ' refers to the woman's individual questionnaire; 'HHS' refers to the household schedule
3. 'More reliable' refers to the data source with less evidence of age misreporting according to the age ratio analysis (see Appendix 2G)
4. The husband's age could only be obtained via the woman's relationship to the household head for women who were themselves the household head or the wife of the household head because for other relationships to the household head it was unclear if(1) the woman's husband was resident in the household at the time of interview; (2) which man was actually her husband.

Appendix 2H

Sample percentage distribution for factors hypothesised to be associated with age at marriage used in graphical chain modelling

MIM code, factor & categories	Region and country[1]												
	Central & South America						West Asia & North Africa			South Asia			
	BRA	COL	DOM	GUT	HAI	PER	EGY	KAZ	UZB	BAN	NEP	PAK	PHI
C Place of residence													
Urban	78.4	72.4	58.0	28.6	45.0	51.9	42.9	47.8	50.8	16.2	12.8	51.3	49.7
Rural	21.6	27.6	41.3	71.2	55.0	48.2	57.0	52.2	49.2	83.6	87.1	48.4	50.1
Missing	0.0	0.0	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.4	0.3
D Married/cohabiting							n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	
Married	72.2	48.0	39.4	58.5	48.7	44.2							89.9
Cohabiting	27.8	52.0	60.6	41.5	51.3	55.8							10.1
Missing	0.0	0.0	0.0	0.0	0.0	0.0							0.0
E Household socio-economic status													
Low	26.4	20.6	23.5	35.8	32.2	26.2	34.9	26.9	24.0	58.8	35.6	33.8	39.7
Average	52.4	34.8	45.7	47.4	42.2	42.1	45.1	29.2	26.8	24.0	36.5	29.1	33.1
High	21.2	44.6	30.8	16.8	25.6	31.7	20.1	43.9	49.2	17.2	27.9	37.2	27.2
Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
F Wife works and earns cash													
No	62.5	64.3	68.7	82.0	59.3	66.2	84.7	63.9	86.3	88.3	92.7	90.2	69.5
Yes	37.5	35.7	31.1	17.7	40.7	33.7	15.3	36.1	13.7	11.6	7.3	9.5	30.2
Missing	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.2	0.3
L Household structure	n.a.	n.a.	n.a.	n.a.		n.a.	n.a.	n.a.	n.a.		n.a.	n.a.	n.a.
Nuclear family					92.8					69.3			
Extended family					7.2					30.7			
Missing					0.0					0.0			

Appendix 2H (continued)

Sample percentage distribution for factors hypothesised to be associated with age at marriage used in graphical chain modelling

MIM code, factor & categories	Region and country [1]												
	Central & South America						West Asia & North Africa			South Asia			
	BRA	COL	DOM	GUT	HAI	PER	EGY	KAZ	UZB	BAN	NEP	PAK	PHI
M Consanguinity	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.		n.a.	n.a.	n.a.	n.a.		n.a.
No							57.7					39.5	
First cousins							24.6					48.2	
Other relation							17.8					12.3	
Missing							0.0					0.0	
O Polygamous husband	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.			n.a.
No											96.5	95.6	
Wife is rank 1											]	]	
Wife is rank 2+											] 3.2	] 4.0	
Missing											0.3	0.4	

MIM code, factor & categories	Region and country [1]																
	East, Central & South Africa										West Africa						
	CAF	COM	KEN	MAW	MOZ	RWA	TAN	UGA	ZAM	ZIM	BEN	BUF	CAM	COT	GHA	MLI	NIG
C Place of residence																	
Urban	37.7	26.7	21.8	30.7	27.3	17.3	26.3	31.9	35.8	37.6	32.5	35.4	57.1	49.1	29.7	31.8	40.8
Rural	62.1	73.3	78.2	69.3	72.1	82.5	72.9	68.0	63.7	62.3	67.0	64.5	42.7	50.9	70.3	68.0	59.2
Missing	0.1	0.0	0.0	0.0	0.5	0.1	0.8	0.1	0.4	0.1	0.6	0.1	0.1	0.0	0.0	0.2	0.1
D Married/cohabiting									n.a.	n.a.		n.a.		n.a.			n.a.
Married	16.5	97.8	94.3	96.2	20.3	52.4	88.9	88.0			89.0		89.0		87.9	97.8	
Cohabiting	83.5	2.2	5.7	3.8	79.7	47.6	11.1	12.0			11.0		11.0		12.1	2.2	
Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			0.0		0.0		0.0	0.0	

Appendix 2H (continued)

Sample percentage distribution for factors hypothesised to be associated with age at marriage used in graphical chain modelling

MIM code, factor & categories	Region and country[1]																
	East, Central & South Africa										West Africa						
	CAF	COM	KEN	MAW	MOZ	RWA	TAN	UGA	ZAM	ZIM	BEN	BUF	CAM	COT	GHA	MLI	NIG
E Household socio-economic status																	
Low	39.4	55.2	29.6	36.8	46.5	48.8	33.0	45.9	40.3	52.1	30.4	32.6	22.2	21.2	35.2	29.7	40.9
Average	43.0	34.1	56.7	39.7	35.7	34.4	37.1	35.8	30.6	20.3	45.1	42.6	41.0	43.6	32.7	49.2	30.9
High	17.6	10.8	13.7	23.5	17.8	16.8	29.9	18.2	29.1	27.6	24.5	24.8	36.9	35.2	32.1	21.1	28.2
Missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
F Wife works and earns cash														n.a.			
No	35.9	71.9	56.2	78.0	91.4	37.4	50.0	68.5	65.3	54.6	20.0	53.4	64.5		37.3	65.0	70.3
Yes	64.0	27.7	43.1	21.9	8.1	62.5	50.0	31.5	34.6	45.3	79.8	46.3	35.5		62.7	34.9	29.5
Missing	0.1	0.4	0.7	0.1	0.5	0.1	0.0	0.0	0.0	0.1	0.2	0.2	0.0		0.0	0.0	0.2
L Household structure						n.a.									n.a.		
Nuclear family	66.3	78.3	92.8	96.6	72.8		77.5	93.1	78.6	94.9	70.5	87.8	94.6	90.8		84.5	92.9
Extended family	33.7	21.7	7.2	3.4	27.2		22.5	6.9	21.4	5.1	29.5	12.2	5.4	9.2		15.5	7.1
Missing	0.0	0.0	0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		0.0	0.0
M Consanguinity		n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
No	90.5																
First cousins	]																
Other relation	] 9.5																
O Polygamous husband																	
No	77.9	82.9	87.9	89.6	79.7	94.5	80.4	79.5	90.2	84.0	59.5	62.5	73.9	71.4	83.0	71.2	78.8
Wife is rank 1	8.5	3.8	]	3.0	7.1	0.8	4.7	7.3	4.4	]	13.5	8.5	4.4	5.7	6.1	6.2	4.8
Wife is rank 2+	13.4	12.9	] 11.3	6.9	11.6	4.4	12.4	13.2	4.8	] 16.0	27.0	28.8	21.4	22.9	10.9	22.6	16.3
Missing	0.0	0.4	0.8	0.6	1.5	0.3	2.5	0.0	0.6	0.0	0.0	0.2	0.3	0.1	0.0	0.0	0.1

Notes for Appendix 2H

1. See page xv for country codes.
2. See over for how the household socio-economic status factor was derived

Appendix 2H (continued)

Deriving an indicator of household socio-economic status

The DHS surveys do not ask about socio-economic status prior to marriage, nor do they obtain any summary indicators that refer to the time of the interview. However, questions are asked in the DHS surveys about the condition of the dwelling and the ownership of a number of household amenities. Responses to these variables can be assigned a score to reflect its desirability as a tentative proxy for socio-economic status. For example, a household with electricity is considered to have a higher level of socio-economic status relative to a household without electricity. A number of other studies have proxied household socio-economic status in this way (e.g. Ohadike, 1978; Isiugo-Abanihe, 1994; Sivaram *et al*, 1995; Mahran *et al*, 1995; Jejeebhoy, 1996).

The following six dichotomous variables are used to derive the index for Chapter 6. A score of one represents a more desirable feature:

Water piped inside/outside dwelling:
- 0: no
- 1: yes

Dwelling has flush toilet to a sewer
- 0: no
- 1: yes

Dwelling has electricity
- 0: no
- 1: yes

Dwelling has a radio:
- 0: no
- 1: yes

Dwelling has a television:
- 0: no
- 1: yes

Dwelling has a fridge:
- 0: no
- 1: yes

Scores range from zero to six. Household socio-economic status can be further categorised to create a variable with three categories, which is more useful for analysis using MIM 3.1. These three categories can be considered as roughly corresponding to low, average and high socio-economic status. Since the scores are seldom evenly distributed, the classification of the three socio-economic statuses must be considered with respect to each country's distribution of socio-economic status scores. This is therefore a country-specific variable.

The percentage distribution of each country's sample with low, average and high socio-economic status according to this measure is shown overleaf.

<table>
<tr><th rowspan="2">Country by region</th><th colspan="7">Household socio-economic status score</th></tr>
<tr><th>0</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th></tr>
<tr><td>Central & South America</td><td colspan="7"></td></tr>
<tr><td>Brazil</td><td colspan="4">28.4</td><td colspan="2">49.7</td><td>22.0</td></tr>
<tr><td>Colombia</td><td colspan="4">20.6</td><td colspan="2">34.8</td><td>44.6</td></tr>
<tr><td>Dominican Republic</td><td colspan="2">17.5</td><td colspan="2">46.0</td><td colspan="3">36.5</td></tr>
<tr><td>Guatemala</td><td colspan="2">28.7</td><td colspan="3">45.1</td><td colspan="2">26.2</td></tr>
<tr><td>Haiti</td><td>33.6</td><td colspan="2">42.2</td><td colspan="4">24.2</td></tr>
<tr><td>Peru</td><td colspan="2">20.8</td><td colspan="3">41.8</td><td colspan="2">37.4</td></tr>
<tr><td>West Asia & North Africa</td><td colspan="7"></td></tr>
<tr><td>Egypt</td><td colspan="4">32.4</td><td colspan="2">45.0</td><td>22.6</td></tr>
<tr><td>Kazakhstan</td><td colspan="4">25.5</td><td>31.1</td><td colspan="2">43.4</td></tr>
<tr><td>Uzbekistan</td><td colspan="4">31.8</td><td>29.4</td><td colspan="2">38.9</td></tr>
<tr><td>South Asia</td><td colspan="7"></td></tr>
<tr><td>Bangladesh</td><td>58.8</td><td>24.0</td><td colspan="5">17.2</td></tr>
<tr><td>Nepal</td><td>38.4</td><td>35.7</td><td colspan="5">25.9</td></tr>
<tr><td>Pakistan</td><td colspan="2">49.2</td><td colspan="2">25.3</td><td colspan="3">25.6</td></tr>
<tr><td>Philippines</td><td colspan="2">39.7</td><td colspan="2">33.1</td><td colspan="3">27.2</td></tr>
<tr><td>Central, East & South Africa</td><td colspan="7"></td></tr>
<tr><td>Central African Republic</td><td>38.6</td><td>43.3</td><td colspan="5">18.1</td></tr>
<tr><td>Comoros</td><td colspan="2">55.2</td><td colspan="2">34.1</td><td colspan="3">10.8</td></tr>
<tr><td>Kenya</td><td>28.9</td><td colspan="2">55.5</td><td colspan="4">15.6</td></tr>
<tr><td>Malawi</td><td>45.0</td><td>39.0</td><td colspan="3">16.0</td><td colspan="2">n.a.</td></tr>
<tr><td>Mozambique</td><td>50.0</td><td>35.9</td><td colspan="5">14.1</td></tr>
<tr><td>Rwanda</td><td>53.7</td><td>35.4</td><td colspan="5">10.9</td></tr>
<tr><td>Tanzania</td><td>35.7</td><td>36.8</td><td colspan="5">27.5</td></tr>
<tr><td>Uganda</td><td>54.7</td><td>35.3</td><td colspan="5">10.0</td></tr>
<tr><td>Zambia</td><td>36.6</td><td>28.4</td><td colspan="5">35.0</td></tr>
<tr><td>Zimbabwe</td><td colspan="2">48.5</td><td colspan="2">20.8</td><td colspan="3">30.8</td></tr>
<tr><td>West Africa</td><td colspan="7"></td></tr>
<tr><td>Benin</td><td>29.4</td><td>43.4</td><td colspan="5">27.1</td></tr>
<tr><td>Burkina Faso</td><td>39.9</td><td>47.1</td><td colspan="5">12.9</td></tr>
<tr><td>Cameroon</td><td>27.7</td><td colspan="2">45.2</td><td colspan="4">27.1</td></tr>
<tr><td>Côte d'Ivoire</td><td>22.1</td><td colspan="2">44.6</td><td colspan="4">33.3</td></tr>
<tr><td>Ghana</td><td>35.2</td><td>32.7</td><td colspan="5">32.1</td></tr>
<tr><td>Mali</td><td>31.3</td><td>49.8</td><td colspan="5">18.9</td></tr>
<tr><td>Niger</td><td>52.8</td><td>33.3</td><td colspan="5">13.8</td></tr>
</table>

Key:

Low	(white)
Average	(light grey)
High	(dark grey)

Appendix 2I

Sample percentage distribution of factors with country-specific categories: Wife's education and husband's education

MIM code, factor & categories	Region and country [1]												
	Central & South America						West Asia & North Africa			South Asia			
	BRA	COL	DOM	GUT	HAI	PER	EGY	KAZ	UZB	BAN	NEP	PHI	PAK
G Wife's highest level of education													
None	]	]	]	28.7	37.6	]	34.2	]	]	47.6	65.0	]	69.5
Primary	] 36.3	] 36.5	] 53.3	55.7	39.6	] 36.2	15.5	]	]	30.6	17.3	] 29.6	10.5
Secondary	57.7	52.2	32.8	]	]	42.4	42.1	] 77.8	] 86.3	]	]	40.3	]
Higher	6.0	11.2	13.9	] 15.6	] 22.8	21.4	8.2	22.2	13.7	] 21.6	] 17.6	30.0	] 20.0
Don't know/missing			0.0	0.0		0.0		0.0	0.0	0.0	0.0	0.0	0.0
H Husband's highest level of education													
None	]	]	]	17.5	22.8	]	17.9	]	]	43.9	28.8	]	38.6
Primary	] 39.9	] 37.0	] 49.5	60.5	36.5	] 26.2	23.6	]	]	25.1	22.0	] 34.1	14.8
Secondary	51.2	48.1	27.9	]	]	48.0	43.5	] 81.2	] 78.7	]	36.9	38.3	]
Higher	6.4	13.2	17.0	] 19.9	] 28.1	25.6	15.0	18.5	21.3	] 30.7	12.4	27.2	] 46.5
Don't know/missing			5.6	2.2	12.6	0.2	0.1	0.3	0.0	0.3	0.0	0.5	0.2

Appendix 2I (continued)

Sample percentage distribution of factors with country-specific categories: Wife's education and husband's education

MIM code, factor & categories	Region and country [1]																
	East, Central & South Africa										West Africa						
	CAR	COM	KEN	MAW	MOZ	RWA	TAN	UGA	ZAM	ZIM	BEN	BUF	CAM	COT	GHA	MLI	NIG
G Wife's highest level of education																	
None	53.1	46.0	7.4	32.2	31.1	30.4	19.9	19.3	13.0	5.6	72.5	77.7	29.8	65.5	47.3	81.0	77.1
Primary	34.1	32.1	61.8	60.3	63.5	57.4	72.9	60.3	62.0	41.3	19.5	]	37.4	22.6	44.4	13.1	14.8
Secondary	]	]	]	]	]	]	]	]	]	]	]	]	]	]	]	]	]
Higher	] 12.8	] 21.9	] 30.8	] 7.5	] 5.4	] 12.2	] 6.9	] 20.5	] 25.0	] 53.1	] 8.0	] 22.3	] 32.7	] 11.9	] 8.3	] 5.9	] 8.1
Don't know/missing	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H Husband's highest level of education																	
None	21.8	46.0	5.7	13.1	15.9	26.2	13.8	7.5	6.4	4.4	50.7	73.8	25.0	51.4	35.0	73.2	76.5
Primary	40.6	17.1	49.1	66.3	58.1	60.0	72.3	55.8	45.4	35.1	21.8	]	34.5	19.6	45.9	11.3	9.6
Secondary	]	]	]	]	]	]	]	]	]	]	]	]	]	]	]	]	]
Higher	] 32.0	] 33.7	] 44.1	] 20.1	] 12.9	] 12.9	] 13.2	] 32.2	] 47.1	] 59.6	] 19.2	] 21.6	] 37.7	] 26.6	] 17.3	] 11.4	] 11.9
Don't know/missing	5.7	3.2	1.1	0.5	13.1	0.9	0.7	4.4	1.1	0.8	8.2	4.7	2.8	2.5	1.8	4.1	2.0

Note for Appendix 2I

1. See page xv for country codes.

Appendix 2I (continued)

Sample percentage distribution of factors with country-specific categories: ethnicity (where data are available)

Region and country	Ethnic group	Percentage distribution
Central & South America		
Brazil		
	White	41.1
	Non-white	58.9
Guatemala		
	Indian	43.7
	Mixed race	56.3
West Asia & North Africa		
Uzbekistan		
	Uzbekistan	80.5
	Else	19.5
South Asia		
Philippines		
	Tagalog	20.0
	Cebuano	30.2
	Ilocano	12.5
	Ilonggo	10.2
	Else	27.1
East, Central & South Africa		
Rwanda		
	Hutu	90.4
	Tutsi	9.2
	Missing	0.4
Zimbabwe		
	Black	97.6
	Else	2.4
West Africa		
Cameroon		
	Cameroonian	95.4
	Other African	4.4
	Missing	0.1
Niger		
	Haoussa	52.7
	Djerma	21.6
	Touareg Bella	9.2
	Else	16.6

Note for Appendix 2I
The ethnic groups are as defined in each country's DHS survey

Appendix 2I

Sample percentage distribution of factors with country-specific categories: religion (where data are available)

Region and country	Religion	Percentage distribution
Central & South America		
Brazil		
	Catholic	79.5
	Protestant	13.5
	Other (including none)	7.0
Guatemala		
	Catholic	55.7
	Protestant	26.6
	Other (mainly none)	17.7
Haiti		
	Catholic	52.0
	Other (mainly Protestant Christian)	48.0
West Asia & North Africa		
Egypt		
	Muslim	94.3
	Christian	5.7
Kazakhstan		
	Muslim	65.8
	Christian	21.3
	'Not religious'	10.8
	Not answered	2.1
Uzbekistan		
	Muslim	94.5
	Else	5.5
South Asia		
Bangladesh		
	Muslim	86.6
	Christian (& 0.5% of 'other')	13.4
Nepal		
	Hindu	87.6
	Else (incl. Buddhist, Muslim, Christian, & other)	12.4
Philippines		
	Roman Catholic	81.2
	Else	18.8

Appendix 2I (continued)

Percentage distribution of religion (where data are available)

Region and country	Religion	Percentage distribution
East, Central & South Africa		
Central African Republic		
	Catholic	34.2
	Protestant	53.5
	Muslim	10.6
	Animist	1.4
	Missing	0.3
Kenya		
	Roman Catholic	30.3
	Protestant	61.7
	Else	8.1
Mozambique		
	Roman Catholic	26.7
	Protestant	29.0
	Muslim	15.1
	Other	6.9
	None	21.2
	Missing	1.1
Rwanda		
	Roman Catholic	62.5
	Protestant	21.9
	7^{th} Day Adventist	11.0
	Else	4.6
Uganda		
	Roman Catholic	40.1
	Protestant	41.3
	Else	18.5
Zambia		
	Roman Catholic	21.8
	Protestant	76.6
	Else	1.6
Zimbabwe		
	Traditional	7.1
	Spiritual	37.1
	Christian	48.8
	Other	7.0
	Missing	0.1

Appendix 2I (continued)

Percentage distribution of religion (where data are available)

Region and country	Religion	Percentage distribution
West Africa		
Benin		
	Traditional	21.4
	Muslim	22.1
	Christian	43.3
	None	12.9
	Missing	0.3
Burkina Faso		
	Muslim	59.0
	Roman Catholic	29.4
	Traditional	10.7
	Missing	1.0
Cameroon		
	Roman Catholic	37.7
	Protestant	31.6
	Muslim	21.7
	None	8.8
	Missing	0.1
Côte d'Ivoire		
	Roman Catholic	16.6
	Protestant	12.8
	Muslim	48.1
	Traditional	3.9
	None	18.6
Ghana		
	None	15.6
	Protestant	45.3
	Roman Catholic	15.9
	Muslim	16.1
	Traditional	7.0
	Missing	0.2
Mali		
	Muslim	90.0
	Christian	3.0
	Animist	5.3
	Other	1.7
Niger		
	Muslim	97.6
	Other	2.4

Appendix 2J

Sample sizes used for the graphical modelling in MIM, by region and country

Region and country [1]	All couples		Monogamous couples only	
	Number [2] of women eligible [3] for this study	Reduced sample size [2,4] (% of total eligible)	Number [2] of women eligible [3] for this study	Reduced sample size [2,4] (% of total eligible)
Central & South America				
Brazil	2717	1994 (73.4)	2717	1994 (73.4)
Colombia	2523	1723 (68.3)	2523	1723 (68.3)
Dominican Republic	1772	1772 (100.0)	1772	1772 (100.0)
Guatemala	3026	1616 (53.4)	3026	1616 (53.4)
Haiti	540	540 (100.0)	540	540 (100.0)
Peru	7307	1909 (26.1)	7307	1909 (26.1)
West Asia & North Africa				
Egypt	5351	1902 (35.5)	5351	1902 (35.5)
Kazakhstan	948	948 (100.0)	948	948 (100.0)
Uzbekistan	1361	1361 (100.0)	1361	1361 (100.0)
South Asia				
Bangladesh	2580	2417 (93.7)	2580	2417 (93.7)
Nepal	2996	2130 (71.1)	2883	2057 (71.3)
Pakistan	2440	2440 (100.0)	2333	2333 (100.0)
Philippines	3694	1965 (53.2)	3694	1965 (53.2)
Central, East & South Africa				
Central African Republic	1373	1373 (100.0)	1070	1070 (100.0)
Comoros	502	502 (100.0)	416	416 (100.0)
Kenya	1229	1229 (100.0)	1080	1080 (100.0)
Malawi	1065	1065 (100.0)	954	954 (100.0)
Mozambique	2236	2014 (90.1)	1777	1777 (100.0)
Rwanda	1396	1396 (100.0)	1319	1319 (100.0)
Tanzania	2148	2128 (99.1)	1728	1728 (100.0)
Uganda	1802	1802 (100.0)	1433	1433 (100.0)
Zambia	2113	2113 (100.0)	1906	1906 (100.0)
Zimbabwe	877	877 (100.0)	737	737 (100.0)
West Africa				
Benin	1626	1617 (99.4)	968	968 (100.0)
Burkina Faso	1642	1642 (100.0)	1027	1027 (100.0)
Cameroon	681	681 (100.0)	503	503 (100.0)
Côte d'Ivoire	1215	1215 (100.0)	867	867 (100.0)
Ghana	660	660 (100.0)	548	548 (100.0)
Mali	2629	2112 (80.3)	1872	1510 (80.7)
Niger	1141	1141 (100.0)	899	899 (100.0)

Notes for Appendix 2J

1. Countries are listed alphabetically within each region
2. Unweighted sample size
3. See section 2.3.2 in Chapter 2 for details of selection criteria
4. Reduced sample was randomly selected

Appendix 2K

Questions used to derive three indices of women's autonomy: Comparing those available from the 1988 Egyptian DHS and used by Kishor (1995) with those available from 1995 Egyptian DHS

Customary autonomy index	
As used by Kishor (1995) from the 1988 Egyptian DHS	**As used in Chapter 6 from the 1995 Egyptian DHS (& variable names)**
"Who should have the last word on the following – the husband, the wife, both, or someone else?"	*"Who has the final say in your family on the following – you or your husband, both you and your husband, or someone else?"*
Having another child	Having another child (wsq304c)
Children's education	Children's education (wsq304d)
Children's marriage plans	Children's marriage plans (wsq304e)
(Not available)	Medical attention for the children (wsq304g)
Use of family planning methods	Use of family planning methods (wsq304h)

Non-customary autonomy index	
As used by Kishor (1995) from the 1988 Egyptian DHS	**As used in Chapter 7 from the 1995 Egyptian DHS (& variable names)**
"Who should have the last word on the following – the husband, the wife, both, or someone else?"	*"Who has the final say in your family on the following – you or your husband, both you and your husband, or someone else?"*
Visits to friends or relatives	Visits to friends and family (wsq304a)
Household budget	Household budget (wsq304b)
Lending or borrowing	(Not available)
	"Now I would like to get your opinion on some aspects of family life. Please tell me if you agree or disagree with each statement:"
(Not available)	A wife's place is not only in the home (wsq305b)
"In general, if a wife disagrees with her husband should she keep quiet or speak up?"	If a wife disagrees with her husband, she should speak-up (wsq305h)
(Not available)	*"Do you know how much your family's approximate total income from all sources is?"* (wsq605)
"Do you think a wife respects a husband more if he insists she accepts his opinion on everything or if he listens to and accepts her opinion?"	(Not available)

Appendix 2K (continued)

Questions used to derive three indices of women's autonomy: Comparing those available from the 1988 Egyptian DHS and used by Kishor (1995) with those available from 1995 Egyptian DHS

Realised autonomy index	
As used by Kishor (1995) from the 1988 Egyptian DHS	**As used in Chapter 7 from the 1995 Egyptian DHS (& variable names)**
"In your home, does your point of view carry the same weight as your husband's, less weight than his point of view, or isn't it taken into account at all?"	(Not available)
"Do you go out with your husband to purchase major household items/clothing?"	(Not available)
"Does your husband allow you to go out alone or with your children to....?"	*"Are you usually allowed to go to the following places on your own, only with children, only with another adult, or not at all?"*
Buy household things	Local market to buy things (wsq314b)
(Not available)	Local health centre/doctors (wsq314c)
(Not available)	Neighbourhood for recreation (wsq314d)
Visit relatives or friends	Home of friends/relatives (wsq314e)
	"Who mainly decides how your family's income is spent?" (wsq606)

Appendix 3A

Pearson correlation matrix of the 12 hypothesised-indicators of women's status used for the factor analysis

Hypothesised indicator of women's status	Hypothesised indicator of women's status (A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)
(A) Adult illiteracy	-											
(B) Secondary school enrolment	-.85**	-										
(C) Female economic activity	-.35**	.32**	-									
(D) Female household headship	.50**	.48**	.58**	-								
(E) Total fertility rate	.86**	-.80**	.48**	.52**	-							
(F) Contraceptive use	-.80**	.76**	.40**	.45**	.88**	-						
(G) Maternal mortality	.69**	-.67**	-.12	-.25*	.61**	-.63**	-					
(H) Extent of safe childbirth	-.82**	.82**	.31**	.42**	-.78**	.75**	-.78**	-				
(I) Life expectancy at birth	-.87**	.87**	.25*	.39**	-.86**	.86**	-.73**	.90**	-			
(J) Representation of women in parliamentary assembly	-.41**	.38**	.41**	.61**	-.44**	.38**	-.24*	.34**	.35**	-		
(K) Representation of women in ministerial posts	-.34**	.34**	.27**	.55**	-.34**	.29**	-.16	.30**	.31**	.71**	-	
(L) Proportion of the population who are Muslim	.62**	-.47**	-.55**	-.59**	.63**	-.48**	.26*	-.50**	-.43**	-.33**	-.33**	-

Notes for Appendix 3A
** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Appendix 4A

Cross-national variations in the prevalence & intensity of polygamy, by region

South Asia

Country	No. of co-wives	Rank within union:				All wives	
		1st wife	2nd wife	3rd + wife	Don't know	%	n
Nepal	0	99.0	Not applicable			96.4	2891
	1	1.0	93.3	0.0	11.1	3.3	99
	2+	0.1	6.7	100.0	0.0	0.3	10
	Don't know	0.0	0.0	0.0	88.9	0.3	8
	Column %	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	3008
	Row %	*97.1*	*2.5*	*0.1*	*0.3*	*100.0*	3008
Pakistan	0	99.5	Not applicable			96.3	2383
	1	0.5	100.0	0.0	0.0	3.2	79
	2+	0.0	0.0	100.0	0.0	0.2	4
	Don't know	0.0	0.0	0.0	100.0	0.3	8
	Column %	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	2474
	Row %	*96.8*	*2.8*	*0.1*	*0.3*	*100.0*	2474

Central, East and South Africa

Country	No. of co-wives	Rank within union:				All wives	
		1st wife	2nd wife	3rd + wife	Don't know	%	n
C.A.R.	0	90.4	Not applicable			78.1	1079
	1	8.7	91.3	0.0	0.0	18.2	251
	2+	0.8	8.7	100.0	0.0	3.5	49
	Don't know	0.0	0.0	0.0	100.0	0.2	3
	Column %	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	1382
	Row %	*86.3*	*11.7*	*1.8*	*0.2*	*100.0*	1382
Comoros	0	95.6	Not applicable			82.9	417
	1	3.0	82.0	0.0	0.0	12.5	64
	2+	1.4	18.0	100.0	100.0	4.2	21
	Don't know	0.0	0.0	0.0	0.0	0.0	0
	Column %	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	502
	Row %	*86.7*	*12.2*	*0.8*	*0.4*	*100.0*	502
Kenya	0	Rank not asked				86.9	1104
	1					8.5	108
	2+					3.7	46
	Don't know					0.9	11
	Column %					*100.0*	*1269*

Appendix 4A (continued)

Central, East and South Africa (continued)

Country	No. of co-wives	Rank within union: 1st wife	2nd wife	3rd + wife	Don't know	All wives: %	n
Malawi	0	97.0	Not applicable			90.4	898
	1	1.6	68.9	0.0	0.0	5.7	57
	2+	1.4	31.1	100.0	0.0	3.8	38
	Don't know	0.0	0.0	0.0	100.0	0.4	4
	Column %	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	997
	Row %	*92.8*	*6.1*	*0.6*	*0.4*	*100.0*	997
Mozambique	0	91.3	Not applicable			79.3	1753
	1	5.0	54.1	0.0	2.9	9.4	208
	2+	3.7	45.9	100.0	41.2	10.4	231
	Don't know	0.0	0.0	0.0	55.9	0.9	20
	Column %	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	2212
	Row %	*86.8*	*9.4*	*2.3*	*1.5*	*100.0*	2212
Rwanda	0	99.1	Not applicable			94.6	1322
	1	0.7	93.9	0.0	0.0	4.1	57
	2+	0.2	6.1	100.0	0.0	1.2	17
	Don't know	0.0	0.0	0.0	100.0	0.1	4
	Column %	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	1400
	Row %	*95.4*	*3.6*	*0.9*	*0.2*	*100.0*	1400
Tanzania	0	94.6	Not applicable			80.0	1769
	1	4.5	80.4	0.0	0.0	11.8	261
	2+	1.1	19.6	100.0	0.0	7.7	170
	Don't know	0.0	0.0	0.0	100.0	0.5	11
	Column %	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	2211
	Row %	*84.7*	*9.9*	*4.9*	*0.5*	*100.0*	2211
Uganda	0	90.6	Not applicable			79.8	1484
	1	6.5	78.0	0.0	0.0	14.1	263
	2+	2.9	22.0	100.0	0.0	6.1	113
	Don't know	0.0	0.0	0.0	0.0	0.0	0
	Column %	*100.0*	*100.0*	*100.0*	*0.0*	*100.0*	1860
	Row %	*88.0*	*10.8*	*1.2*	*0.0*	*100.0*	1860
Zambia	0	95.7	Not applicable			90.7	1928
	1	3.7	81.8	0.0	0.0	6.9	146
	2+	0.6	18.2	100.0	0.0	1.8	39
	Don't know	0.0	0.0	0.0	100.0	0.6	13
	Column %	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	2126
	Row %	*94.7*	*4.1*	*0.5*	*0.6*	*100.0*	2126
Zimbabwe	0	Rank not asked				85.2	786
	1					12.3	114
	2+					2.5	23
	Don't know					0.0	0
	Column %					*100.0*	*923*

Appendix 4A (continued)

West Africa

Country	No. of co-wives	Rank within union: 1st wife	2nd wife	3rd + wife	Don't know	All wives %	n
Benin	0	82.3	Not applicable			60.5	978
	1	16.1	85.0	0.0	0.0	29.4	474
	2+	1.7	14.7	100.0	0.0	10.1	163
	Don't know	0.0	0.3	0.0	0.0	0.1	1
	Column %	*100.0*	*100.0*	*100.0*	*0.0*	*100.0*	1616
	Row %	*73.6*	*20.7*	*5.8*	*0.0*	*100.0*	1616
Burkina Faso	0	85.9	Not applicable			57.3	1008
	1	11.0	79.4	0.0	25.0	25.8	455
	2+	3.2	20.6	100.0	0.0	16.7	292
	Don't know	0.0	0.0	0.0	75.0	0.2	3
	Column %	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	1758
	Row %	*66.8*	*23.2*	*9.8*	*0.2*	*100.0*	1758
Cameroon	0	93.5	Not applicable			69.9	477
	1	3.5	60.0	0.0	0.0	13.8	94
	2+	2.9	40.0	100.0	0.0	15.9	109
	Don't know	0.0	0.0	0.0	100.0	0.4	3
	Column %	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	683
	Row %	*74.8*	*18.3*	*6.5*	*0.4*	*100.0*	683
Côte d'Ivoire	0	92.7	Not applicable			71.5	860
	1	6.9	87.3	0.0	0.0	20.7	249
	2+	0.4	12.7	100.0	0.0	7.7	93
	Don't know	0.0	0.0	0.0	100.0	0.1	1
	Column %	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	1203
	Row %	*77.2*	*17.6*	*5.1*	*0.1*	*100.0*	1203
Ghana	0	93.2	Not applicable			83.0	548
	1	4.6	53.3	0.0	0.0	8.9	59
	2+	2.2	46.7	100.0	0.0	8.0	53
	Don't know	0.0	0.0	0.0	0.0	0.0	0
	Column %	*100.0*	*100.0*	*100.0*	*0.0*	*100.0*	660
	Row %	*89.1*	*9.1*	*1.8*	*0.0*	*100.0*	660
Mali	0	91.7	Not applicable			70.9	1886
	1	7.9	93.2	0.0	0.0	24.6	654
	2+	0.4	6.8	100.0	0.0	4.5	120
	Don't know	0.0	0.0	0.0	100.0	0.0	1
	Column %	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	2661
	Row %	*77.3*	*19.8*	*2.8*	*0.0*	*100.0*	2661
Niger	0	94.5	Not applicable			79.6	916
	1	5.4	92.5	0.0	0.0	17.4	200
	2+	0.1	7.5	100.0	0.0	2.9	33
	Don't know	0.0	0.0	0.0	100.0	0.1	1
	Column %	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	1150
	Row %	*84.2*	*14.0*	*1.7*	*0.1*	*100.0*	1150

References

Aaby, P., K. Ariyoshi, M. Buckner, H. Jensen, N. Berry, A. Wilkins, D. Richard, O. Larsen, F. Dias, M. Melbye and H. Whittle (1996) "Age of wife as a major determinant of male-to-female transmission of HIV-2 infection: a community study from rural West Africa," *AIDS*, 10: 1585-1590.

Adepoju, A. (1983) "Patterns of migration by sex," Chapter 3 in C. Oppong ed. *Female and male in West Africa.* London: Allen and Unwin.

Adepoju, A. and C. Oppong (1994) "Introduction" in A. Adepoju and C. Oppong eds. *Gender, work and population in sub-Saharan Africa.* London: James Currey.

Adepoju, A. (1994) "Chapter 2" in A. Adepoju and C. Oppong eds. *Gender, work and population in sub-Saharan Africa.* London: James Currey.

Adepoju, A. (1994) "Women, work and fertility in Swaziland" Chapter 10 in A. Adepoju and C. Oppong eds. *Gender, work and population in sub-Saharan Africa.* London: James Currey.

Adepoju, A. and C. Oppong eds. (1994) *Gender, work and population in sub-Saharan Africa.* London: James Currey.

Agarwala, S. N. (1957) "The age at marriage in India," *Population Index*, 23 (2): 96-107.

Akers, D. S. (1967) "On measuring the marriage squeeze," *Demography*, 4 (2): 907-924.

Assaad, M. B. (1980) "Female circumcision in Egypt: Social implications, current research, and prospects for change," *Studies in Family Planning*, 11 (1): 3-16.

Atkinson, M. P. and B. L. Glass (1985) "Marital age heterogamy and homogamy, 1900 to 1980," *Journal of Marriage and the Family*, 43 (3): 685-691.

Balk, D. (1994) "Individual and community aspects of women's status and fertility in rural Bangladesh," *Population Studies*, 48 (1): 21-45.

Balk, D. (1997) "Defying gender norms in rural Bangladesh: A social demographic analysis," *Population Studies*, 51 (3): 153-172.

Barbieri, M. and V. Hertrich (1999) *The effect of conjugal closeness on fertility behaviour in sub-Saharan Africa.* Paper presented in session 55 'Marriage Systems and Fertility Change in Sub-Saharan Africa' at the 1999 Annual Meeting of the Population Association of America held in New York, March 25-27, 1999.

Bartiaux, F. (1994) "The declining birth rate and the marriage market: Italy 1930-1950", *Population*, 49 (1): 91-117.

Bascom, W. R. and M. J. Herskovits eds. (1959) *Continuity and change in African cultures.* Chicago: Chicago Press.

Becker, G. S. (1973) "A Theory of Marriage, Part I," *Journal of Political Economy*, 81 (4): 813-846.

Becker, G. S. (1974) "A Theory of Marriage, Part II," *Journal of Political Economy*, 82 (2), Part 2: S11-S26.

Berardo, F. M, Appel, J., Berardo, D. H. (1993) "Age Dissimilar Marriages: Review and Assessment", *Journal of Aging Studies*, 7, (1): 93-106.

Bisilliat, J. (1983) "The feminine sphere in the institutions of the Songhay-Zarna," chapter 6 in C. Oppong ed. *Female and male in West Africa*. London: Allen and Unwin.

Bittles, A. H., J. C. Grant and S. A. Shami (1993) "Consanguinity as a determinant of reproductive behaviour and mortality in Pakistan," *International Journal of Epidemiology*, 22 (3): 463-467

Blake, J. (1974) "The changing status of women in developed countries," *Scientific American*, 231 (3): 137-147.

Blanc, A. K., and N. Rutenberg (1990) "Assessment of the quality of data on age at first sexual intercourse, age at first marriage, and age at first birth in the Demographic and Health Surveys," pp. 37-79 in Macro International ed. *An Assessment of DHS-I Data Quality*. Demographic and Health Surveys Methodological Reports No. 1. Columbia, MD, USA: Macro International.

Bloom, D. E., and P. H. Reddy (1986) "Age patterns of women at marriage, cohabitation, and first birth in India," *Demography*, 23 (4): 509-523.

Bloom, S. S., D. Wypij, and M. Das Gupta (2001) "Dimensions of women's autonomy and the influence on maternal health care utilization in a north Indian city," *Demography*, 38 (1): 67-78.

Bogue, D. J. (1969) "Marriage and marital dissolution" chapter 17 in D. J. Bogue (ed.) *The Principles of Demography*. New York: John Wiley and Sons, Inc.

Bongaarts, J. (1978) "A framework for analyzing the proximate determinants of fertility," *Population and Development Review*, 4 (1): 105-132.

Bongaarts, J., O. Frank and R. J. Lesthaeghe (1984) "The proximate determinants of fertility in sub-Saharan Africa," *Population and Development Review*, 10 (3): 511-537.

Boserup, E. (1970) *Women's role in economic development*. London: Allen and Unwin.

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

Bozon, M. and Heran, F. (1989) "Finding a spouse: a survey of how French couples meet". *Population: An English Selection*, 44: 90-121.

Bozon, M. (1991) "Women and the age gap between spouses: An accepted domination?," *Population: An English Selection*, 3: 113-148.

Brass, W., A. J. Coale, P. Demeny, D. F. Heisel, F. Lorimer, A. Romaniuk, and E. van de Walle (1968) *The Demography of Tropical Africa*. Princeton, New Jersey: Princeton University Press.

Brown, J. (1981) "Polygyny and family planning in sub-Saharan Africa," *Studies in Family Planning*, 12 (89): 322-326.

Buckley, C. J. (1996) "Gender, age and the marriage market: Evidence on marriage in late adulthood in Russia," *Journal of Cross-Cultural Gerontology*, 11 (3): 255-267.

Bulatao, R. A. and R. D. Lee eds. (1983) *Determinants of fertility in developing countries*. Vol. 2. San Diego, California: Academic Press.

Bumpass, L.L. and J. A. Sweet (1972) "Differentials in marital instability," *American Sociological Review*, 37 (6): 754-766.

Cain, K. C., R. A. Kronmal, and A. S. Kosinski (1992) "Analyzing the relationship between change in a risk factor and risk of disease," *Statistics in Medicine*, 11: 783-797.

Cain, M. T. (1984). *Women's Status and Fertility in Developing Countries: Son Preference and Economic Security*. World Bank Staff working paper No. 682, Population and Development Series No. 7. Washington, D.C.: The World Bank.

Cain, M. T. (1993) "Patriarchal structure," chapter 2 in N. Federici, K. O. Mason and S. Sogner eds. *Women's position and demographic change*. Oxford: Clarendon.

Caldwell, J. C. (1963) "Fertility decline and female chance of marriage in Malaya," *Population Studies*, 17 (1): 20-32.

Caldwell, J. and P. Caldwell (1993) "Women's position and child mortality and morbidity in less developed countries," pp.122-139 in N. Federici, K. O. Mason and S. Sogner eds. *Women's position and demographic change*. Oxford: Clarendon.

Campbell, M. J. and M. J. Gardner (1988) "Calculating confidence intervals for some non-parametric analyses," *British Medicine Journal*, 296: 1454-1456.

Caraël, M. (1995) "Sexual behaviour," chapter 4 in J. Cleland and B. Ferry eds. *Sexual behaviour and AIDS in the developing world*. London: Taylor and Francis.

Carrier, N.H. (1959) "A note on the measurement of digital preference in age recordings," *Journal of the Institute of Actuaries*, 85: 71-85.

Casterline, J. B., L. Williams, and P. McDonald (1986) "The age difference between spouses: Variations among developing countries," *Population Studies*, 40: 353-374.

Central Intelligence Agency (1999) *The World Factbook 1999*. http://www.cia.gov/cia/publications/factbook/ Washington D.C: Central Intelligence Agency, U.S. Government.

Chamie, J. (1986) "Polygyny among Arabs," *Population Studies*, 40 (1): 55-66.

Cherlin, A. and A. Chamratrithirong (1988) "Variations in marriage patterns in central Thailand," *Demography*, 25 (3): 337-353.

Child, D. (1970) *The essentials of factor analysis*. London: Holt, Rinehart and Winston

Chowdhury, F. I. and F. Trovato (1994) "The role and status of women and the timing of marriage in five Asian countries," *Journal of Comparative Family Studies*, 25 (2): 143-157.

Cleland, J., C. Scott and D. Whitelegge eds. (1987) *The World Fertility Survey: An assessment*. International Statistical Institute, London: Oxford University Press.

Cleland, J. and B. Ferry eds. (1995) *Sexual behaviour and AIDS in the developing world*. London: Taylor and Francis.

Cleveland, D. A. (1991) "Migration in West Africa: A Savannah village perspective," *Africa*, 61 (2): 222-246.

Cox, P. R. (1970a) "Sex differences in age at marriage," *Journal of Biosocial Science*, Suppl. 2: 73-84.

Cox, P. R. (1970b) "International variations in the relative ages of brides and grooms," *Journal of Biosocial Science*, 2: 111-121.

Cox, D. R. and N. Wermuth (1996) *Multivariate dependencies: Models, analysis and interpretation*. London: Chapman and Hall.

Croft, T. (2000) Personal communication.

De Silva, W. I. (1998) "Puerperal morbidity: a neglected area of maternal health in Sri Lanka," *Social Biology*, 45 (3-4): 223-245.

Dinan, C. (1983) "Sugar daddies and gold-diggers: the white-collar single women in Accra," chapter 25 in C. Oppong ed. *Female and male in West Africa*. London: Allen and Unwin.

Dixon, R. B. (1971) "Explaining cross-cultural variations in age at marriage and proportions never marrying", *Population Studies*, 25 (2): 215-233.

Dorjahn, V. R. (1959) "The factor of polygyny in African demography," chapter 5 in W. R. Bascom and M. J. Herskovits eds. *Continuity and change in African cultures*. Chicago: Chicago Press.

Dyson, T. and M. Moore (1983) "On kinship structure, female autonomy, and demographic behaviour in India," *Population and Development Review*, 9 (1): 35-60.

Easterlin, R. A. (1973) "Relative economic status and the American fertility swing", pp. 170-223 in E. B. Sheldon, ed. *Family Economic Behavior: Problems and Prospects*. Philadelphia: J. B. Lippincott.

Edwards, D. (1995) *Introduction to Graphical Modelling*. Springer Texts in Statistics. New York: Springer-Verlag.

Ehrenberg, A.S.C. (1975) *Data reduction: Analysing and interpreting statistical data*. Chichester: Wiley.

Ekpere, J. A., F. P. A. Oyedipe, and R. O. Adegboye (1978) "Family role differentiation within the Kwara Nomadic Fulani," pp. 55-67 in C. Oppong, G. Adaba, M. Bekombo-Priso, and J. Mogey eds. *Marriage, fertility and parenthood in West Africa*. Canberra: Australian National University Press.

El-Zanaty, F., E. M. Hussein, G. A. Shawky, A. A. Way, and S. Kishor (1996) *Egypt Demographic and Health Survey 1995*. Calverton, Maryland, USA and National Population Council, Egypt: Macro International Inc.

Epstein, E. and R. Guttman (1983) "Mate selection in man: Evidence, theory, and outcome," *Social Biology*, 31 (3-4): 243-278.

Ewbank, D.C. (1981) *Age misreporting and age-selective under-enumeration: Sources, patterns and consequences for demographic analysis*. Committee on Population and Demography Report no. 4. Washington D.C.: National Academy Press.

Fapohunda, E. R. (1983) "Female and male work profiles," chapter 2 in C. Oppong ed. *Female and male in West Africa*. London: Allen and Unwin.

Federici, N., K. O. Mason, and S. Sogner eds. (1993) *Women's position and demographic change*. Oxford: Clarendon.

Feng, W. and Y. Quanhe (1996) "Age at marriage and the first birth interval: The emerging change in sexual behaviour among young couples in China," *Population and Development Review*, 22 (2): 299-320.

Fortes, M. (1978) "Family, marriage and fertility in West Africa," pp.17-54 in C. Oppong, G. Adaba, M. Bekombo-Priso, and J. Mogey eds. *Marriage, fertility and parenthood in West Africa*. Canberra: Australian National University Press

Foster, A. D. and N. U. Khan (1994) *Equilibrating the marriage market in a rapidly growing population: Evidence from rural Bangladesh*. Economics Department and Population Studies Center Working Paper, University of Pennsylvania.

Fox, A. J., L. Bulusu, and L. Kinlen (1979) "Mortality and age differences in marriage," *Journal of Biosocial Science*, 11: 117-131.

Fricke, T. E., S. H. Syed, and P. C. Smith (1986) "Rural Punjabi social organisation and marriage timing strategies in Pakistan," *Demography*, 23 (4): 489-508.

Gage, A. J. (1995) "An assessment of data on age at first union, first birth, and first sexual intercourse for phase II of the Demographic and Health Surveys program". Demographic and Health Surveys Occasional Papers No. 4. Calverton, Maryland: Macro International.

Garenne, M. and E. van de Walle (1989) "Polygyny and fertility among the Sereer of Senegal," *Population Studies*, 43 (2): 267-283.

Goldman, N., C. Westoff, and C. Hammerslough (1984) "Demography of the marriage market in the United States," *Population Index*, 50 (1): 5-25.

Goldman, N., S. O. Rustein, and S. Singh eds. (1985) *Assessment of the quality of data in 41 WFS Surveys: A comparative approach*. WFS Comparative Studies No. 44. Voorburg, Netherlands: International Statistical Institute.

Goldman, N. and A. Pebley (1989) "The demography of polygyny in sub-Saharan Africa," chapter 5 in R. J. Lesthaeghe ed. *Reproduction and social origin in sub-Saharan Africa*. London: University of California Press.

Govindasamy, P. and A. Malhotra (1996) "Women's position and family planning in Egypt," *Studies in Family Planning*, 27 (6): 328-340.

Greene, M. E. and V. Rao (1995) "The marriage squeeze and the rise in informal marriage in Brazil," *Social Biology*, 42 (1-2): 65-82.

Hagan, G. P. (1983) "Marriage, divorce and polygyny in Winneba," chapter 13 in C. Oppong ed. *Female and male in West Africa*. London: Allen and Unwin.

Hajnal, J. (1953) "Age at marriage and proportions marrying," *Population Studies*, 7 (2): 111-136.

Hallouda, A. M., S. Farid, and S. H. Cochrane eds. (1988) *Egypt: Demographic responses to modernisation*. Cairo, Egypt: Central Agency for Public Mobilisation and Statistics.

Hampshire, K. and S. Randall (2000) "Pastoralists, agropastoralists and migrants: Interactions between fertility and mobility in northern Burkina Faso," *Population Studies*, 54 (3): 247-261.

Harris, J .A. (1912) "Assortive mating in man," *Population Science Monthly*, 80: 476-492.

Henry, L. (1972) "Nuptiality," *Theoretical Population Biology*, 3: 135-152.

Hollingshead, A. B. (1950) "Cultural factors in the selection of marriage mates," *American Sociological Review*, 15 (5): 619-627.

Hollingshead, A. B. (1951) "Age relationships and marriage," *American Sociological Review*, 16 (4): 492-499.

Inter-Parliamentary Union (1997) *Men and women in politics: Democracy still in the making*. Geneva: Inter-Parliamentary Union.

Isiugo-Abanihe, U. C. (1994) *Nuptiality patterns, sexual activity, and fertility in Nigeria*. Demographic and Health Surveys Working Paper No. 16. Calverton, Maryland: Macro International Inc.

Islam, M. N. and A. U. Ahmed (1998) "Age at first marriage and its determinants in Bangladesh," *Asia-Pacific Population Journal*, 13 (2): 73-92.

Jejeebhoy, S. J. (1991) "Women's status and fertility: Successive cross-sectional evidence from Tamil Nadu, India, 1970-1980," *Studies in Family Planning*, 22 (4): 217-229.

Jejeebhoy, S. J. (1996) *Women's education, autonomy and reproductive behaviour: Assessing what we have learned.* Paper prepared for the conference on 'The Status of women and demographic change: Assessing what we have learned' held at the East-West Center, Honolulu, Hawaii, December 18-20, 1995.

Jensen, A. R. (1978) "Genetic and behavioral effect of non-random mating," pp. 51-105 in C. E. Noble, R. T. Osborne and N. Weyle eds. *Human variations: Biogenetics of age, race and sex.* New York: Academic Press.

Johnson, J. J. and W. Turnbull (1995) "The women's conference: Where aspirations and realities met," *Family Planning Perspectives*, 27 (6): 254-258.

Jones, G. W. (1980) "Trends in marriage and divorce in Peninsular Malaysia," *Population Studies*, 34 (2): 272-292.

Kalmijn, M. (1998) "Intermarriage and homogamy: Causes, patterns and trends," *Annual Review of Sociology*, 24: 395-421.

Kishor, S. (1995) *Autonomy and Egyptian women: Findings from the 1988 Egypt Demographic and Health Survey.* Demographic and Health Surveys Occasional Paper No. 2. Calverton, Maryland: Macro International Inc.

Kishor, S. and K. Neitzel (1996) *The status of women: Indicators for twenty-five countries.* Demographic and Health Surveys Comparative Studies No. 21. Calverton, Maryland: Macro International Inc.

Kritz, M. M. and D. T. Gurak (1989) "Women's status, education and family formation in sub-Saharan Africa," *International Family Planning Perspectives*, 15 (3): 100-105.

Lampard, R. (1993) "Availability of marriage partners in England and Wales: A comparison of three measures," *Journal of Biosocial Science*, 25: 333-350.

Lê, T. N. and V. K. Verma (1997) *An analysis of sample designs and sampling errors of the Demographic and Health Surveys*. Demographic and Health Surveys Analytical Reports No.3. Calverton, Maryland: Macro International Inc.

Lesthaeghe, R. J., G. Kaufmann, and D. Meekers (1989) "The nuptiality regimes in sub-Saharan Africa," chapter 6 in R. J. Lesthaeghe ed. *Reproduction and social organisation in sub-Saharan Africa.* London: University of California Press.

Lindenbaum, S. (1981) "Implications for women of changing marriage transactions in Bangladesh," *Studies in Family Planning*, 12 (11): 394-401.

Linton, R. (1942) "Age and sex categories," *American Sociological Review*, 7: 589-603.

Macro International (2001) http://www.measuredhs.com/

Magnani, R., J. T. Bertrand, B. Makini, and S. McDonald (1995) "Men, marriage and fatherhood in Kinshasa, Zaire," *International Family Planning Perspectives*, 21 (1): 19-25.

Mahran, M., F. H. El-Zanaty, and A. A. Way (1995) *Perspectives on Fertility and Family Planning in Egypt.* Cairo, Egypt: National Population Council and Calverton, Maryland: Macro International Inc.

Makina, P. and A. Jensen eds. (1995) *Women's position and demographic change in sub-Saharan Africa.* Liège: International Union for the Scientific Study of Population.

Malhotra, A. (1991) "Gender and changing generational relations: Spouse choice in Indonesia," *Demography*, 28 (4): 549-570.

Malthus, T. R. (1798) *An Essay on the Principle of Population.* Oxford: Oxford University Press.

Manly, B.F.J. (1986) *Multivariate Statistical Methods: A primer.* London: Chapman and Hall.

Martin, T. C. (1995) "Women's education and fertility: Results from 26 Demographic and Health Surveys," *Studies in Family Planning*, 26 (4): 187-202.

Mason, K. O. (1985) *The status of women. A review of its relationship to fertility and mortality.* New York: The Rockefeller Foundation.

Mason, K. O. (1986) "The status of women: Conceptual and measurement issues in demographic studies," *Sociological Forum*, 1 (2): 284-300.

Mason, K. O. (1987) "Differences between women's and men's reproductive goals in developing countries," *Population and Development Review*, 13 (4): 611-638.

Mason, K.O. (1993) "Women's position and demographic change," pp. 19-43 in N. Federici, K. O. Mason and S. Sogner eds. *Women's position and demographic change.* Oxford: Clarendon.

Mason, K. (1995) *Is the situation of women in Asia improving or deteriorating?* Asia-Pacific Population Research Report No. 6. Honolulu, Hawaii: East-West Center Program on Population.

Mazur, R. and M. Mhloyi (1994) "Women's work and fertility in Zimbabwe: Ending under-development with change," chapter 9 in A. Adepoju and C. Oppong eds. *Gender, work and population in sub-Saharan Africa.* London: James Currey.

McCarthy, J. (1982) *Differentials in age at first marriage.* WFS Comparative Studies. Cross national summaries No. 19. Voorburg, Netherlands: International Statistical Institute.

McFarland, D. D. (1970) "Effects of group size on the availability of marriage partners," *Demography*, 7 (4): 411-415.

Meekers, D. (1992) "The process of marriage in African societies: A multiple indicator approach," *Population and Development Review*, 18 (1): 61-78.

Meekers, D. (1996) "Combining ethnographic and survey methods: A study of the nuptiality patterns of the Shona of Zimbabwe," *Journal of Comparative Family Studies*, 25 (3): 313-328.

Mhloyi, M. (1994) *Status of women, population and development.* International Union for the Scientific Study of Population distinguished lecture series on population and development. Liege, Belgium: International Conference on Population and Development.

Mohamed, W. N., I. Diamond, and P. W. F. Smith (1998) "The determinants or infant mortality in Malaysia: a graphical chain modelling approach," *Journal of the Royal Statistical Society Series A*, 161 (3): 349-366.

Morah, B. C. (1985) *Evaluation of the Nigeria Fertility Survey 1981-2.* WFS Scientific Report No. 80. Voorburg, Netherlands: International Statistical Institute.

Nagi, M.H., E. G. Stockwell, and L. M. Snavley (1973) "Digit preference and avoidance in the age statistics of some recent African censuses: Some patterns and correlates," *International Statistical Review*, 41 (2): 165-174.

Nawar, L. M. (1988) "Household structure and age at marriage in Abbasai, Cairo," chapter 4 in A. M. Hallouda, S. Farid and S. H. Cochrane eds. *Egypt: Demographic responses to modernisation*. Cairo, Egypt: Central Agency for Public Mobilisation and Statistics.

Neugarten, B., J. W. Moore, and J. Lowe (1965) "Age norms, age constraints, and adult socialisation," *American Journal of Sociology*, 70 (6): 710-717.

Ní Bhrolcháin, M. (1992) "Age difference asymmetry and a two-sex perspective," *European Journal of Population*, 8: 23-45.

Ní Bhrolcháin, M. (1997) *The age difference between partners: a matter of female choice?* Paper prepared for the meeting on 'Age –Between Nature and Culture' organised by the International Union for the Scientific Study of Population Working Group on Biology, Culture and Population, Rostock, 29-31 May 1997.

Ní Bhrolcháin, M. (2000) *Flexibility in the marriage market*. Department of Social Statistics Working Paper No. 00-02. Department of Social Statistics, University of Southampton.

Niraula, B. B. and S. P. Morgan (1996) "Marriage formation, post-marital contact with natal kin and autonomy of women: Evidence from two Nepali settings," *Population Studies*, 50 (1): 35-50.

Norušis, M. J. (1992) *SPSS for Windows. Professional Statistics. Release 5.* Chicago: SPSS Inc.

Obermeyer, C. M. (1992) "Islam, women, and politics: The demography of Arab countries," *Population and Development Review*, 18 (1): 33-60.

Ogbuagu, S. C. (1996) *Rural dimensions of gender-population concerns*. Paper prepared for a workshop on 'Gender, Population and Development' organised by UNFPA in association with UNIFEM, New York, 30 September - 3 October, 1996.

Ohadike, P. O. (1978) A case study of the determinants of family household size in African development. Pp. 318-398 in Oppong, C., G. Adaba, M. Bekombo-Priso, and J. Mogey eds. *Marriage, fertility and parenthood in West Africa*. Canberra: Australian National University Press.

Oppenheimer, V. K. (1988) "A theory of marriage timing," *American Journal of Sociology*, 94 (3): 563-591.

Oppenheimer, V. K., M. Kalmijn, and N. Lim (1997) "Men's career development and marriage timing during a period of rising inequality," *Demography*, 34 (3): 311-330.

Oppong, C., G. Adaba, M. Bekombo-Priso, and J. Mogey eds. (1978) *Marriage, fertility and parenthood in West Africa*. Canberra: Australian National University Press.

Oppong, C. (1983) *Female and male in West Africa*. London: Allen and Unwin.

Oppong, C. (1995) "Some roles of women: what do we know? Conceptual and methodological issues in sub-Saharan Africa," pp. 363-388 in P. Makina and A. Jensen eds. *Women's position and demographic change in sub-Saharan Africa*. Liège: International Union for the Scientific Study of Population.

Park, C. B. and N. H. Cho (1995) "Consequences of son preference in a low-fertility society: Imbalance of the sex ratio at birth in Korea," *Population and Development Review*, 21 (1): 59-84.

Parrado, E. M. and M. Tienda (1997) "Women's roles and family formation in Venezuela: New forms of consensual unions?" *Social Biology*, 44 (1-2): 1-24.

Parsons, T. (1940) "An analytical approach to the theory of social stratification," *American Journal of Sociology*, 45 (6): 841-862.

Parsons, T. (1942) "Age and sex in the social structure of the United States," *American Sociological Review*, 7 (5): 604-616.

Pinnelli, A. (1993) "The condition of women and the health and mortality of infants and children," pp. 162-189 in N. Federici, K. O. Mason and S. Sogner eds. *Women's position and demographic change.* Oxford: Clarendon.

Plewis, I. (1985) *Analysing change: Measurement and explanations using longitudinal data.* Chichester: Wiley

Pollard, A.H., Yusuf, E., and Pollard, G.N. (1990) *Demographic Techniques.* Third edition, Sydney: Pergamon Press.

Powers, M.G. and Salvo, J. J. (1982) "Fertility and childcare arrangements as mechanisms of status articulation," *Journal of Marriage and the Family,* 44 (1): 21-34.

Presser, H. B. (1975) "Age differences between spouses: Trends, patterns, and social implications.," *American Behavioural Scientist*, 19: 190-205.

Rashad, H. (2000) "Demographic transition in Arab countries: A new perspective," *Journal of Population Research*, 17 (1): 83-101.

Remez, L. (1998) "In Turkey, women's fertility is linked to education, employment, and freedom to choose a husband," *International Family Planning Perspectives*, 24 (2): 97-98.

Riley, N. E. (1997) "Gender, power and population change," *Population Bulletin*, 52 (1): 2-48.

Rugh, A. B. (1984) *Family in contemporary Egypt.* New York: Syracuse University Press.

Rutstein, S. O and G. T. Bicego (1990) *Assessment of the quality of data used to ascertain eligibility and age in the Demographic and Health Surveys.* Demographic and Health Surveys Methodological Reports No.1. Colombia, Maryland: Institute for Resource Development/Macro Systems, Inc.

Sachedina, Z. (1990) "Islam, procreation and the law," *International Family Planning Perspectives*, 16 (3): 107-111.

Salaff, J. W. (1976) "The status of unmarried Hong Kong women and the social factors contributing to their delayed marriage," *Population Studies*, 30 (3): 391-412.

Sall, J., Lehman, A., and Creighton, L. (2001) *JMP Start Statistics, A Guide to Statistics and Data Analysis Using JMP and JMP IN Software.* 2nd edition. Belmont, California: Duxbury Press.

Schildkrout, E. (1983) "Dependence and autonomy: the economic activities of secluded Hausa women in Kano," chapter 7 in C. Oppong ed. *Female and male in West Africa.* London: Allen and Unwin.

Schmitt, R. C. (1965) "Demographic correlates of inter-racial marriages in Hawaii," *Demography*, 2: 466.

Schoen, R. (1983) "Measuring the tightness of a marriage squeeze," *Demography*, 20 (1): 61-78.

Schuler, S. R., S. M. Hashemi, and A. P. Riley (1997) "The influence of women's changing roles and status in Bangladesh's fertility transition: Evidence from the study of credit programs and contraceptive use," *World Development*, 25 (4): 563-575.

Seidenberg, R. (1972) "Older women and younger men," *Sexual Behavior*, April: 9-17.

Shah, N. M. and E. Q. Bulatao (1981) "Purdah and family planning in Pakistan," *International Family Planning Perspectives*, 7 (1): 32-37.

Sheela, J. and N. Audinarayana (1997) "Work status of women and age at marriage in Coimbatore city: A multivariate analysis," *Indian Journal of Social Work*, 58 (1): 78-89.

Sheldon, E. B. ed. (1973) *Family Economic Behavior: Problems and Prospects*. Philadelphia: J. B. Lippincott.

Shryock, H. and Siegel, J. (1976) *The methods and materials of demography*. Volume 1. Washington D.C.: U.S. Department of Commerce, Bureau of the Census. US Government Printing Office.

Simmons, R. and A. M. Young (1995) *Family Planning Programs and other interventions to assist women: Their impact on demographic change and on the status of women*. Paper prepared for the conference on 'The Status of women and demographic change: Assessing what we have learned', held at the East-West Center, Honolulu, Hawaii, December 18-20, 1995.

Simmons, R. and Young, A.M. (1996) Family Planning Programs and other interventions to assist women: Their impact on demographic change and on the status of women. Honolulu, Hawaii: East-West Center Program on Population.

Singh, S. (1985) "Assessment of nuptiality data," chapter 3 in N. Goldman, S. O. Rustein, and S. Singh eds. *Assessment of the quality of data in 41 WFS Surveys: A comparative approach*. WFS Comparative Studies No. 44. Voorburg, Netherlands: International Statistical Institute.

Singh, S. (1987) "Evaluation of data quality," chapter 24 in J. Cleland, C. Scott and D. Whitelegge eds. *The World Fertility Survey: An assessment*. International Statistical Institute, London: Oxford University Press.

Singh, M. (1992) "Changes in age at marriage of women in rural north India," *Journal of Biosocial Science*, 24: 123-130.

Singh, S. and R. Samara (1996) "Early marriage among women in developing countries," *International Family Planning Perspectives*, 22 (4): 148-157.

Sivaram, M., J. Richard, and P. S. S. Rao (1995) "Early marriage among rural and urban females of South Asia," *Journal of Biosocial Science*, 27: 325-331.

Smith, P. C. (1983) "The impact of age at marriage and proportions marrying on fertility," chapter 13 in R. A. Bulatao and R. D. Lee eds. *Determinants of fertility in developing countries*. Vol. 2. San Diego, California: Academic Press.

Smith, H. L. (1989) "Integrating theory and research on the institutional determinants of fertility," *Demography*, 26 (2): 171-183.

Sokona, O. and J. B. Casterline (1988) "Socio-economic differentials in age at marriage," chapter 3 in A. M. Hallouda, S. Farid and H. Cochrane eds. *Egypt: Demographic responses to modernisation*. Cairo, Egypt: Central Agency for Public Mobilisation and Statistics.

Stoeckel, J. and J. Chowdhury (1984) "Spouse age difference and fertility in rural Bangladesh: Implications for raising the age at marriage of females," *Biology and Society*, 1 (3): 139-143.

Timaeus, I. and W. Graham (1989) "Labour circulation, marriage and fertility in southern Africa," pp. 365-400 in R. J. Lesthaeghe ed. *Reproduction and social organisation in sub-Saharan Africa.* London, University of California Press.

Timaeus, I. and A. Reynar (1998) "Polygynists and their wives in sub-Saharan Africa: An analysis of five Demographic and Health Surveys," *Population Studies*, 52: 145-162.

United Nations (1984) *Compiling social indicators on the situation of women.* Studies in Methods Series F, No.32. New York: Department of International Economic and Social Affairs Statistical Office and International research and Training Institute for the Advancement of women, United Nations.

United Nations (1987a) "Women's employment and fertility," chapter 9 in United Nations ed. *Fertility behaviour in the context of development: Evidence from the World Fertility Surveys.* Department of International Economic and Social Affairs Population Studies No. 100. New York: United Nations.

United Nations (1987b) *Fertility behaviour in the context of development: Evidence from the World Fertility Surveys.* Department of International Economic and Social Affairs Population Studies No. 100. New York: United Nations.

United Nations (1990) *Patterns of first marriage: Timing and Prevalence.* New York, United Nations

United Nations (1994) *Women's Indicators and Statistics Database Volume 3.* New York, United Nations

United Nations (1999) *Human Development Report 1999.* New York: Oxford University Press.

Van de Walle, E. (1968) "Marriage in African censuses and inquiries," chapter 5 in W. Brass, A. J. Coale, P. Demeny, D. F. Heisel, F. Lorimer, A. Romaniuk, and E. van de Walle eds. *The Demography of Tropical Africa.* Princeton, New Jersey: Princeton University Press.

Van de Walle, F. and E. van de Walle (1993) "Urban women's autonomy and natural fertility in the Sahel region of Africa," pp. 61-80 in N. Federici, K. O. Mason and S. Sogner eds. *Women's position and demographic change.* Oxford: Clarendon.

Veevers, J. E. (1988) "The 'real' marriage squeeze," *Sociological Perspectives*, 31 (2): 169-189.

Vellenga, D. D. (1983) "Who is a wife? Legal processes of heterosexual conflicts in Ghana," chapter 9 in C. Oppong ed. *Female and male in West Africa.* London: Allen and Unwin.

Villarreal, M. (1998) *Adolescent fertility: Socio-cultural issues and programming implications.* Sustainable Development Department, Food and Agriculture Organization of the United Nations, http://www.fao.org/WAICENT/FAOINFO/SUSTDEV/WPdirect/wpre0072.htm.

Vlassoff, C. (1992) "Progress and stagnation: Changes in fertility and women's position in an Indian village," *Population Studies*, 46 (2): 195-212.

Walsh, B. M. (1972) "Trends in age at marriage in post-war Ireland," *Demography*, 9 (2): 187-202.

Warren, B. D. (1966) "A multiple variable approach to the assortative mating phenomenon," *Eugenics Quarterly*, 13: 285-290.

Weekes, R. (1984) *Muslim Peoples: A World ethnographic survey.* London: Aldwych Press

Wheeler, R. H. and B. G. Gunter (1987) "Change in spouse age difference at marriage: A challenge to traditional family and sex roles?" *Sociological Quarterly*, .3 (3): 411-421.

Winch, R. F. and S. A. Greer (1964) "The uncertain relationship between early marriage and marital stability: A quest for relevant data," *Acta Sociologica*, 8 (1-2): 83-97.

World Health Organisation (1991) *Coverage of maternity care: A listing of information available.* Geneva: World Health Organisation.

Xenos, P. and S. A. Gultiano (1992) *Trends in female and male age at marriage and celibacy in Asia.* Papers of the Program on Population No. 120. Honolulu, Hawaii: East-West Center Program on Population.

Yanez III, N. D., R. A. Kronmal, and L. R. Shemanski (1998) "The effects of measurement error in response variables and tests of association of explanatory variables in change models," *Statistics in Medicine*, 17: 2597-2606.

Zhang, W. (2000) "Dynamics of marriage change in Chinese rural society in transition: a study of a northern Chinese village," *Population Studies*, 54: 57-69.