Social Interactions, Reproductive Externalities, and Fertility Behaviour in Kenya

Sriya Iyer
Faculty of Economics
University of Cambridge

Melvyn Weeks*
Faculty of Economics
University of Cambridge

August 2, 2005

Abstract

This paper examines, theoretically and empirically, the impact of reproductive externalities on fertility behaviour in Kenya. We examine this issue by identifying structural forms of social interaction operating across individuals belonging to different ethnic groups on the number of children ever born. We use the 1998 Demographic and Health Survey data on 5994 women from Kenya to examine whether the 'local' effect of household-level influences, cluster-level residential settlement, and the more 'global' effect of ethnicity, is important over and above an individual's characteristics in order to explain variations in fertility behaviour in this country. In so doing, we conclude that the importance of social interactions for fertility is that it may be necessary to target population policy towards varied ethnic groups in different ways, with relevance for population policy in ethnically pluralistic societies more widely.

JEL Classification: C25, D1, J13, J15, Z13

Keywords: fertility behaviour, strategic complementarities, social interaction, endogenous effects, ethnicity, Kenya .

*corresponding author: email:mw217@econ.cam.ac.uk.

1 Introduction

Both economists and demographers have examined the balance between economic and non-economic factors in orchestrating a fertility transition, as witnessed in historical European populations and in some East Asian economies over relatively short periods of time. Classic studies of fertility such as the Princeton Fertility Project highlighted that the transition to low fertility in historical European populations occurred in a variety of socio-economic and institutional contexts, with a significant role being played by the local social environment (Coale and Watkins (1986), Federici, Mason, and Sogner (1993)). More recent studies have emphasised the role of the social climate and the influence of social interactions on demographic behaviour (Egero and Hammraskjold (1994), Hank (2001). These empirical findings have led many economists naturally to focus on modelling the influence of social interactions on contemporary fertility transitions (Durlauf and Walker (1999); Manski (1993), Manski and Mayshar (2002)) and on the reproductive externalities and coordination failures associated with fertility behaviour per se ((Dasgupta (2000); Kohler (2001)).

Empirical and theoretical analyses of social interactions can be divided into two broad approaches. One approach considers the impact of social interactions within predetermined groups (Akerlof (1997); Brock and Durlauf (2000), Brock and Durlauf (2001)), and emphasises how interactions affect individual and group-level outcomes in a cross-section. A second approach analyses how group formation results from social interactions with particular emphasis on the growth of residential neighbourhoods. This literature has focused on questions of geographic proximity and dynamic group formation (Borjas (1995); Ginther, Haverman, and Wolfe (2000); Conley and Topa (2002)).

Our study is located within the first strand of the literature and examines social interactions in the context of fertility behaviour in Kenya where groups are considered predetermined. We conceive of the fertility of a Kenyan woman to be influenced by a range of factors such as her individual characteristics and the characteristics of the household to which she belongs. We postulate that the social environment impinges on a woman's behaviour through the multiple interactions that she undertakes both at the household level and at the level of other reference groups. In doing so we highlight a crucial point often overlooked in both the theoretical and empirical literature on this issue: namely that there are different levels at which social interactions occur. In doing so we consider the possibility that both *local and global forms of interaction* may coexist and be important, characterised by different mechanisms of social interaction (Horst and Scheinkman (2003a)). We argue that this has profound implications, both theoretically and empirically, for studies of fertility behaviour and social interactions, and for the study of social interactions and economic phenomenon, more widely. The research presented in this paper quantifies the effects of social interactions and group membership on the number of children ever born. In this respect our work highlights the existence of uncertainty as to whether these effects are mediated through household composition, ethnic affiliation, a neighbourhood cluster effect, or some combination thereof. There is thus uncertainty both with respect to the levels at which interactions occur, and the expressions of structural dependence at each level.

Section 2 summarizes the literature on social interactions and fertility behaviour. In section 3 we provide an overview of ethnic groups in Kenya, highlighting the characteristics which are relevant for fertility behaviour. In section 4 we present a model of reproductive externalities with social utility effects. In sections 5 we consider the distinction between multiple levels of social interaction, and in sections 6 and 7 we examine, respectively, the question of identification and endogeneity. In section 8 we examine a number of non-linear models of fertility behaviour, and in section 9 we introduce the data. We discuss our results in section 10, and section 11 concludes.

2 Social interactions and fertility

The conventional economic household demand model of fertility behaviour posits that a couple's fertility is a function of the money costs of children and the opportunity costs of the value of parental time (Becker (1981)). Conditional upon identifying a fixed set of fundamentals as comprised of a vector of characteristics of both the individual and the household, an atomistic model of fertility behaviour simply focusses on the direct affects attributable to a change in these fundamentals. Subsequently, if we then observe that the variance in fertility outcomes is in excess of that which might be accounted for by differences in fundamentals,[1] an extension of the standard fertility model to allow for the existence of social multiplier effects would seem reasonable (see, for example, Becker and Murphy (2000)).

Theoretical work of Dasgupta (2000) and Kohler (2001) has argued that a couple's fertility may be influenced by the level of fertility of all other couples within a society. In the most general sense, this is what is meant by social interaction in a fertility context: it is the public interaction between individuals in a society as they perceive each other and observe each others' fertility behaviour. This consequently alters their social environment, which in turn ultimately influences their private decision-making about fertility. As a consequence, the existence of social interactions may lead to multiple equilibria and coordination failures in demographic decision-making, shown by a high level of variability in outcomes for a given set of fundamentals. Recent economic analyses of fertility behaviour have been much concerned with social interactions (see Montgomery, Kiros, Agyeman, Casterline, Aglobitse, and Hewett (2001), Hank (2001), Nauck (1995), Becker and Murphy (2000), Kohler (2001), and Manski (2000)).

In the social interaction literature[2] and following Manski (1993), three hypotheses are often advanced to explain the observation that individuals belonging to a common reference group tend to behave similarly, even after controlling for a set of observed individual characteristics. In a fertility context, there is an endogenous effect if, ceteris paribus, a woman's children ever born (CEB) tends to vary with

[1] Horst and Scheinkman (2003b) note that any change in the fundamentals exerts direct effects and indirect effects that have the same sign. They comment that when the indirect effects are very significant, the multiplier is very large.

[2] For an excellent discussion of the theoretical literature on economic models of social interactions and the influence of non-market institutions on market institutions see Glaeser and Scheinkman (1999)).

the average CEB of members of her ethnic group or locality, perhaps because she suffers a utility loss from deviations from group behaviour. There is an exogenous or contextual effect if, ceteris paribus, a woman's CEB tends to vary with (or more plausibly inversely with) the average educational attainment of her ethnic group or locality. For example, the existence of educated neighbours may foster positive attitudes towards smaller family size. Finally, there is a correlated effect if, ceteris paribus, women in the same ethnic group or locality tend to have similar CEB because they are, for example, similarly wealthy. The importance of differentiating between these three hypotheses can be seen by examining the respective policy implications. Consider, for example, an intervention to provide free contraception to some members of an ethnic group or a number of neighbourhoods. If there are endogenous effects, then an effective policy of free contraception may both directly reduce the fertility of the recipients but, as their fertility decreases, indirectly reduces the fertility of all other members of that ethnic group or neighbourhood, with a feedback to further fertility reductions by the recipients of the free contraception. Exogenous effects and correlated effects will not, in general, generate this kind of social multiplier effect.

It is possible to conceive of a large number of exogenous effects; in the specific context of fertility behaviour, education plays a particularly prominent role. The effect of social interactions on fertility behaviour operates through education at multiple levels. For example, the potential effects of education on fertility can be decomposed into 2 components. The first is the effect of individual educational attainment on individual-level fertility. In a developing society, this would be dependent on the money-costs of acquiring an education and the opportunity costs of wages foregone. Second, there is also an iconic value of education in the sense that some individuals may aspire to the attributes of other higher educated groups in a population. In some developing societies, the iconic value of an education is very high. For example, sociologists of India often comment on the phenomenon of 'Sanskritization' in which lower-caste groups take on the characteristics and customs of the upper castes in order to gain greater legitimacy and status in the Indian social system. Frequently this manifests itself in the desire to acquire an education, or to continue one (Srinivas (1994)).

It is also important to make the distinction between expected versus observed behaviour in locating the mechanism by which a group impacts upon individual decision making. One of the central issues here is the notion that in large groups, individuals do not observe the behaviour of group members, and therefore the appropriate object is to consider interaction effects as mediated by subjective expectations. However, such an argument does not hold in cases where the size of the group is small, such as interaction effects within extended families of small communities.

3 Overview of ethnic groups in Kenya

Social interactions and channels of message transmission about fertility behaviour are important at the level of ethnicity. Members of different ethnic groups speak the same language, usually adopt similar cultural practices, and with some exceptions in border areas, reside closely in the same, or in contiguous districts (Fapohunda and Poukouta (1997)). In the specific case of Kenya it has been argued that cultural norms may encourage high fertility but also contribute to fertility decline over time (Caldwell and Caldwell (1987); Ascadi, Ascadi, and Bulatao (1990)). We argue that individuals' multiple levels of social interactions reflect their social identities - regional identities, ethnic identities, religious identities, linguistic identities. While we do not explore the question of identity in this paper, we do focus on how local and global interactions that arise from these identities, significantly affect individuals' behaviour. Our analysis is based upon historical and anthropological studies of Kenya with particular focus on the features of Kenya's major ethnic groups that have relevance for fertility behaviour. In this section we discuss the ethnic groups with reference to their population composition, the region of their residence, and three characteristics of these groups which are noteworthy - their residential pattern of settlement, clan organisation and education. These characteristics have implications for the politics of ethnicity and fertility in Kenya.

Population The population of Kenya consists of three main groups - the Africans, Asians and Europeans. Although the last two groups mainly reside in towns, 90% of the African population continues to live in rural areas (see Meck (1971)). The tribal groups of rural Kenya live in clearly defined settlements in the more remote areas. In the Lake Victoria basin, the highlands and the coast, there is a more heterogeneous population structure (Morgan and Shaffer (1966):2; Meck (1971): 24). The tribes can be defined broadly into four groups as classified by their language: the Bantu, the Nilotic, the Nilo-Hamitic, and the Hamitic. Within these four broad language groups, ethnic groups represent an additional sub-division. According to the 1969 *Population Census* there are 42 different ethnic African groups in Kenya (Rep (1970)). The five largest groups, accounting for over 75% of the population, are the Kikuyu (22%), Kamba (11%), Kalenjin (12%), Luhya (14%) and Luo (18%).[3] The Kisii and the Meru both comprise 6% of the total population. Other groups, such as the Maasai and the Somali constitute 15% of the population.[4] A religious breakdown of the ethnic groups show that most Kikuyu are Christians, probably because they were the Kenyan tribe with the closest links with Christian missions historically (Meck (1971): 27). Over 90% of Luhya and Luo are Christians, compared to over 60% of Kamba. Some of the Mijikenda sub-groups, such as the Digo, are Muslim.

Census data from Kenya depict a steady increase in the total population of Kenya since 1948 to the present. The total fertility rate in Kenya has fallen from an average of 6 births per woman in 1948 to approximately 4.5 births today (Ajayi and Kekovole (1998)). However, the most significant declines in fertility have occurred in the last twenty years: the DHS studies conducted in 1989, 1993 and 1998 showed that the TFR dropped from 6.7 in 1989 and 5.4 in 1993 to 4.7 in 1998. This drop in fertility is considered one of the most dramatic recorded anywhere in the developed

[3]According to the most recent Census.

[4]The Asians, Europeans and Arabs make up about 1% of the population.

and developing world (Ajayi and Kekovole (1998): 116). The fertility differences by ethnic group are also very large, as shown in Table 1. For women aged 15-49, the mean CEB varies from 2.5 for the Kikuyu to 3.5 for the Luo. As shown in Table 2, these differences are very pronounced for the oldest cohort of women aged 50-59 for whom mean CEB varies from 5.91 for the Kikuyu to 7.56 for the Luo, and who have completed their fertility. For younger women, we only observe truncated fertility outcomes as they may not, as yet, have completed their family size outcomes. Previous studies of fertility and ethnicity in Kenya have revealed that desired family size is smaller among the Kamba, Kikuyu, Luhya compared to the Luo (Fapohunda and Poukouta (1997)). The Luo exhibit a strong preference for large families and it is reported use contraception much less than other groups (Watkins, Rutenberg, and Green (1995)).

Region Ethnic boundaries in Kenya, to a very large extent, are coterminous with political and administrative boundaries (Fapohunda and Poukouta (1997)). Today the regional breakdown of the different ethnic groups is as follows: the Kalenjin reside in the Rift Valley; the Kikuyu live in the Central region, but have also migrated to Nairobi and the Rift Valley. The Meru/Embu reside in the North and East. The Luhya live in the Western province, but have also migrated to Nairobi and Mombasa. The Luo live in Nyanza, with Kisumu as their capital, but have also migrated to the Rift Valley. The Kamba live in close proximity to Nairobi and exhibit ethnic affiliation to the Kikuyu. The Mijikenda/Swahili live in the Coast province. The Meru and Embu groups neighbour the Kikuyu to the north and east. The Luhya, who live in Kenya's Western province, are a less homogenous group compared to the Kikuyu and Kamba (Were (1967)). Among the Luhya, there is a distinction that needs to be made between those in the north and those in areas such as Kakamega, with high population density. Although the Luo who live in Nyanza depict a rate of population growth almost as high as the Luhya, they are a more urbanised people (Ominde (1968)). For the Kamba, proximity to Nairobi and their close ethnic affiliation to the Kikuyu are particularly significant (Berg-Schlosser (1984). The Mijikenda and Swahili groups who reside in Kenya's Coast province are the most agricultural of all of Kenya ethnic groups.[5]

Residential pattern of settlement One of the most notable features of these ethnic groups is their pattern of residence.[6] It is important to emphasise that the notion of the 'village' in this society is rather more diffuse than in the compact settlements of Europe and elsewhere. Although there are differences between ethnic groups, families are grouped mainly in homesteads that are located in clusters, and which form the basis of community interaction. In some of these clusters, for example, among the Kikuyu there exist groups of households who are not merely

[5] The word 'Mijikenda' means 'nine towns' or 'tribes'. The Mijikenda consist of nine distinct sub-groups - Giriama, Duruma, Digo, Rabai, Chonyi, Kambe, Kauma, Ribe, and Jibana.

[6] A key point to note here is that villages are not defined units of settlement, although their boundaries are well-known locally. These boundaries are usually marked by trees, stones, and so forth (Berg-Schlosser (1984): 139)

coresident, but also related by blood and marriage. In other clusters, there are households that are located in close proximity, but where individuals in these households are not related to other individuals in neighbouring households.

For example, the Kalenjin live mainly in homesteads that are individual family based. Property is usually inherited along paternal lines or male agnates. The smallest unit of territorial composition in Kalenjin society is the '*temenik'* or hamlet, a cluster of homesteads; these are usually grouped into 'villages' of 15 to 60 *temenik*. The Kikuyu live in homesteads on land owned by the family, and surrounded by fields (or '*shamba*'). Traditionally land was owned by many households which constitute the extended family located on the same 'ridge' which is the traditional geographical division[7] (Berg-Schlosser (1984): 50). Supplementing the organisation of Kikuyu society based on kinship, there is also a geographical demarcation. The '*itura*' or village consists of groups of families. This is further subdivided into those living on the same ridge. The pattern of Luo settlement is similar to other communities, although their homesteads usually consist of families who own land, dispersed as a safeguard against climatic conditions. The Kamba are also agriculturists, and their pattern of settlement is similar to the Kikuyu (Middleton and Kershaw (1965)). The traditional land-owning unit is the extended family. In terms of territorial residence, the most basic unit is the *ukambani,* which is a homestead that comprises several extended families. Several of these are grouped to form a *kivalo* within which social interactions, especially marriage, take place. For the Mijikenda, the main form of residence is the *mudzi* or village which consists of groups of agriculturists and fisherman.

Clan organisation These tribal groups are also broadly characterised by three main features that have relevance for social interactions: the importance of the family group, the clan, and the system of age-grading[8] (Meck (1971)). For example, anthropological studies of the agriculturalist Kikuyu argue that among them, fathers exert a less important role for economic decision-making, that relations between the extended family are strong, and that they exhibit a highly evolved sense of ethnic identity which can often override more national concerns (Berg-Schlosser (1984), Ferguson and SrungBoonmee (2003)). Anthropologists comment that among the Kikuyu, the most economically and socially effective unit is the *mbari*: 'a group of families who trace their descent from a common ancestor following the paternal line, often for up to seven or eight generations. In addition to its functions as the most important traditional land-holding unit in Kikuyu society, the *mbari* is an important social reference-group for many Kikuyu and still plays an effective role in many economic and social relationships including the more modern ones.' (Berg-Schlosser (1984): 53). We would therefore expect that the endogenous social interaction effects among this group would emerge as being particularly strong. The Luhya do not have as powerful a clan organisation as some of the other groups

[7] This was altered with the land reforms of 1950 when ownership of land was transfered to individuals.

[8] The system of age-grading is a form of vertical stratification which every member of a tribe goes through during the course of their lives, with each 'age-grade' made distinct from the other with certain rites of passage (Kenyatta (1961):2; Berg-Schlosser (1984): 55).

(Meck (1971): 28). Anthropologists comment that a feature of the Luhya that makes them quite distinct from other groups is the relatively strict accepted norms of behaviour, particularly with respect to marriage and interactions between the sexes (Berg-Schlosser (1984): 114).

Education Initiated first by Jomo Kenyatta, education policy in Kenya has been pursued actively by the government, with remarkable success, especially with increases in primary and secondary school enrolment for girls. However, despite the general increases in the uptake of education and policies such as school fee remission and the development of *harambee* or self-help community schools which foster these, the mean number of years of schooling differs considerably by ethnicity.

The mean number of years of education varies considerably by ethnic group in Kenya, as shown in Table 3. This factor has relevance for social interactions as we might expect the importance of social interactions on fertility to be mediated by the effect of education. Previous studies of Kenyan fertility argue that the Kikuyu and the Kamba show the least preference for large families because they had early access to colonial education (Fapohunda and Poukouta (1997)). The Kikuyu are the best educated, compared to other ethnic groups in Kenya (Berg-Schlosser (1984): 60; De Wilde (1967): 39). We can explain this development historically - many Kikuyu worked as wage labourers in European-owned plantations and attended schools. As a community, they viewed education as the means to progress and this led them to acquire positions of responsibility in the colonial administration. In the post-Independence period, the Kikuyu continued to dominate and consolidate their position, politically and economically, relative to the other groups (Ferguson and SrungBoonmee (2003)). The Luhya populations, are also highly literate groups. Most Kamba are enrolled in formal schooling. Among the Kalenjin, levels of education are, in general, less than other groups. The Mijikenda and Swahili groups also depict low levels of education compared to other groups in this population.

Politics of ethnicity Recent evidence has highlighted that differences in the economic performance of ethnic groups can be traced to their development historically and to economic decisions made by these groups (Ferguson and SrungBoonmee (2003)). Ethnic identity in Kenya has always been very strong historically, and it continues to be a very potent force in Kenya's politics even today. For example, the Kalenjin, who are a heterogenous ethnic group and who live primarily in Kenya's Rift Valley province (Were (1967)) are closest to Nairobi geographically, and they are politically and economically Kenya's dominant ethnic group (Kenyatta (1966)). Their location in the eastern Rift Valley allows them to enjoy a level of political proximity that, in part, determines their (relatively) superior economic status. Agricultural innovation first arose among the Kikuyu who also absorbed land reform more readily than other groups (Meck (1971): 27). Historically, the Kikuyu were one of the first ethnic groups in Kenya to absorb European-style capitalism in the form of wage labour and participation in the monetary economy, so in contrast for example to the Maasai whom they neighbour, the Kikuyu's subsequent economic performance has been much better (Ferguson and SrungBoonmee (2003)). [9] In the past the Luo were less responsive to social and economic change compared to other tribes (Meck (1971)). For example, land reform met with a great deal of resistance among this group as it was believed to conflict with religious beliefs. But since 2001, the Luo have increasingly been incorporated into the government. In more recent times, relations between the ethnic groups has been closely tied with the development of Kenyan politics, particularly since the introduction of multiparty politics in Kenya since 1991 after over twenty years of rule by one party (Throup (2001)). The increased strength and importance of the ethnic identity has also led, more worryingly, to ethnic clashes in 1992-93 and in 1997.

So ethnicity - or ethnic group identity - is important for this society. Ethnicity has displayed an immutability which may be accounted for by the historical evolution of these population groups, manifest in the recent political history of this nation.

4 A Model of Fertility with Social Interactions

We consider a population in which there is a representative household h[10] characterised by a location j, a type s - here ethnicity - and a vector of characteristics $\mathbf{x}_{i.}$. We assume that both j and s are fixed for all individuals, with characteristics potentially varying across individuals, type and location. Our point of departure is the standard Beckerian utility function (Becker (1981), Willis (1973))

$$U_i = U(n_i, \mathbf{z}_i, \mathbf{x}_i, \boldsymbol{\varepsilon}_i), \tag{1}$$

where U_i denotes the utility function of household i, n_i represents the number of children, $\mathbf{z}_i$ is all sources of satisfaction to the husband and wife other than those arising from children, and $\mathbf{x}_i$ denotes the vector of socio-demographic characteristics which affect preferences. By including the stochastic term ε_i we allow for imperfect information on the part of the analyst. This is subject to the usual budget constraint:

$$I = n_i p_n + \mathbf{z}_i \mathbf{p}_z, \tag{2}$$

where we assume that total income I is expended on children n_i and on all sources of satisfaction other than those arising from children $\mathbf{z}_i$; p_n and $\mathbf{p}_z$ are the shadow prices of children and other sources of satisfaction respectively.

In extending the above atomistic utility model to incorporate a theory of social interactions we first write

$$V_i = (U(n_i, \mathbf{z}_i, \mathbf{x}_i); S^j(n_i, \mathbf{n}^j_{-i}), S^s(n_i, \mathbf{n}^s_{-i}), \boldsymbol{\varepsilon}_i), \tag{3}$$

[9] For example, following the Land Transfer Programme, Kikuyu land owners in the Central province specialised in the production of coffee which is widely exported by Kenya, earning these farmers resources that ensured their success as a community (Ferguson and SrungBoonmee (2003)).

[10] Note that within the household we consider husband's and wife's preferences to be coincident. In this respect we do not examine intra-household bargaining between the couple in the determination of fertility outcomes. For theoretical simplicity, we assume that the 'household' represents the couple. Empirically this can be translated into a measure of the woman's fertility, measured by the total number of children borne by her.

where V_i, total individual utility, is comprised of a private utility term $U()$, and two social utility terms: $S^j(n_i, \mathbf{n}_{-i}^j)$ represents a social utility term for location j, given by some function of the number of children choices made by all other residents of j; $\mathbf{n}_{-i}^j = (n_1^j, n_2^j, ..., n_{i-1}^j, n_{i+1}^j, ..., n_{M_j}^j)$, where M_j denotes the number of individuals in location j. $S^s(n_i, \mathbf{n}_{-i}^s)$ represents a social utility term for type s, namely the impact of some function of the number of children choices made by all other members of type s, where $\mathbf{n}_{-i}^s = (n_1^s, n_2^s, ..., n_{i-1}^s, n_{i+1}^s, ..., n_{M_s}^s)$. In both cases the extension allows us to capture that component of individual utility attributable to a fertility choice n_i, that is dependent upon the fertility choices of others. The 'others' in question may be those, for example, resident in the individual's household, or others resident in the local community for example, in the individual's street or neighbourhood. Rewriting (3) as additive in individual and social utility terms, we have

$$V_i = U(n_i, \mathbf{z}_i, \mathbf{x}_i) + S^j(n_i, \mathbf{n}_{-i}^j) + S^s(n_i, \mathbf{n}_{-i}^s) + \varepsilon_i. \tag{4}$$

In extending (1) by adding social utility functions $S^j(.)$ and $S^s()$, our intention is to develop a model which facilitates a distinction between *local* and *global* interactions in the sense that individual behaviour may also be determined by the decisions of a larger ethnic group. In referring to a local interaction, individuals have incentives to conform to the behaviour of a small number of appropriately defined neighbours, whose actions they can 'observe'. In contrast a global interaction refers to a situation where individuals face incentives to conform to the 'expected' behaviour of a common reference group (Horst and Scheinkman (2003b)), whose behaviour we cannot observe. What is significant here is that both a local and a global interaction can occur simultaneously. This specification in (3), including more than a single social utility term, allows for *multiple social interactions* - namely the possibility that social interaction is both localised within a specific location, and more diffuse across a larger, geographically diffuse sub-population classified by type. Such a formulation allows us to examine, for example, whether after controlling for social interaction effects accruing through geographical proximity, there is a residual effect due to the desire to conform to a set of behaviours ascribed by ethnicity.

In focussing on that part of total utility attributable to social utility functions $S^j(.)$ and $S^s(.)$, it is obvious from (4) that cross-partial effects will be key objects. Differentiating the Beckerian first-order conditions with respect to social interactions yields the second-order conditions

$$\begin{aligned} \kappa_{il}^j &= \partial^2 S^j(n_i, n_l)/\partial n_i \partial n_l > 0, & (5) \\ \kappa_{il}^s &= \partial^2 S^s(n_i, n_l)/\partial n_i \partial n_l > 0, & (6) \end{aligned}$$

where κ_{il}^j represents a measure of the disutility accruing to i from deviating from the behaviour of l for $i, l \in j$. In examining (5) and (6) we define strategic complementarity as representing the increasing marginal utility of woman i as a direct result of the fertility choices of other women in the group. The precise form of interaction will obviously depend on the functions S^j and S^s.

If we are willing to assume that social utility exhibits strategic complementarities that are totalistic and constant (see Cooper (1988)), then we may impose a number of restrictions. For example, the restriction $\kappa_{il}^j = \kappa^j/2(M_j - 1)$ is consistent with a model of uniform *local* interaction, with equal weights assigned to all members of location j. We note that although such a restriction may be theoretically difficult to justify, it is not possible to identify separate measures κ_{il}^j if there is no information as to the relative location of all i, l within each location. Such a restriction implies that, in the case of type, all individuals are equally influential with respect to fertility decisions; and, in the case of spatially defined clusters, all individuals are located at the centroid.

We consider two forms of social utility which are consistent with totalistic and constant strategic complementarity. At the level of location j these may be written as

$$S_1^j(n_i, \overline{n}_{-i}^j) = \theta_1 n_i \overline{n}_{-i}^j \tag{7}$$

$$S_2^j(n_i, \overline{n}_{-i}^j) = \frac{-\theta_2}{2}(n_i - \overline{n}_{-i}^j)^2. \tag{8}$$

(7) represents a proportional spillover form of dependence in the sense that for $\partial S_1^j(.)/\partial n_i = \theta_1 \overline{n}_{-i}^j$, strategic complementarities are solely dependent upon the mean level of children ever born within location j. (8) represents social utility as a measure of conformism with $\partial S_2(.)/\partial n_i = \theta_2(\overline{n}_{-i}^j - n_i)$. Under (8) there is an incentive to conform to mean fertility behaviour within the group with deviations from the mean penalised more severely, relative to the proportional spillovers case. If behaving like others confers additional status on a woman, she may desire to conform.[11] Note also that although

$$\frac{\partial^2 S_l^j(.)}{\partial n_i \partial \overline{n}_{-i}^j} = \theta_l,\ l = 1, 2,$$

such that strategic complementarities are captured by a single parameter in both cases.

We note that in anticipating subsequent data constraints we have represented the social utility function in a way which assumes that all interactions within a group have equal weight. However, the additional restriction on the social utility terms (as in 7 and 8) that any estimated endogenous interactions are constant over a population is unlikely to hold, and is not required for identification. For example, in the case of effects which stem from a desire to conform to a group fertility norm, such a norm may vary *across* different ethnic groups. In the case of the Kikuyu, which reside in clusters comprised of an extended family, then we have prior information suggesting the (ceteris paribus) likelihood of greater interaction than other groups.

5 Measuring Social Interactions: The Choice of Reference Groups

Given that our data contain no information which points to the existence of specific groups of individuals which define a network within which social interaction takes

[11] There may also be a compelling reason to conform if such status in the community bears economic benefits, e.g. in the allocation of local resources.

place[12] we acknowledge uncertainty as to what constitutes an appropriate reference group with respect to fertility decisions. However, using a number of anthropological sources we have identified that within rural populations in Kenya, fertility decisions are conducted within families which are members of ethnic groups and which reside mainly in homesteads that are located in relatively small clusters. These various allegiances, we postulate, then form the basis of social interaction.

Subsequently we emphasise the role of three reference groups. First, *ethnicity* postulates that the status of women and attitudes towards children may differ substantially across ethnic groups, with differential evaluation of the psychological costs and benefits of bearing children.[13]

Second, we have noted that in the case of Kenya, situated between the individual, the household and the ethnic group, are relatively small *clusters of households* within which physical proximity dictates that individuals directly observe and bear the costs of the decisions of others. In addition we have information which allows us to distinguish between clusters which, for certain ethnic groups, are comprised of households which are blood related, with obvious ramifications for interaction.[14] For example, in the case of the Kikuyu we might expect the extent of such interaction to be on average higher given that clusters of households are generally comprised of groups of individuals which are related by blood and marriage (the *mbari*). In such a situation one might expect to find a stronger normative influence relative to clusters comprised of households, as in the case of the Kalenjin, who live mainly in homesteads that are individual family based.

The third reference group is that within the *household*, and consists of, for example, the influence of extended family on fertility. In studies of South Asia (Hajnal (1982), Iyer (2002)) the household which defines the locus of interaction can be very widely defined, including not only the immediate family, but domestic servants, distant relatives and so forth. The household in Kenya constitutes a well-defined and fixed unit of analysis. There are a number of ways in which interactions at the level of the household might be significant for fertility behaviour. For example, if a woman observes the past fertility behaviour of other female relatives in the household, for example her sisters-in-law, then this might affect her individual decision-making. In the presence of polygynous households, the behaviour of other wives is going to be important, either because it justifies a higher fertility norm, or because there is more help provided with child care. Given that household composition is often crucial to fertility decisions, these intra-household externalities may have consequence for fertility decisions and need to be modelled and taken into account explicitly.

---

[12] For an example where such data exists, see Montgomery, Kiros, Agyeman, Casterline, Aglobitse, and Hewett (2001)

[13] There has been a great deal of research on the effects on fertility of religion and ethnic group membership (Gellner (1981); Sander (1995); Iyer (2002)).

[14] In other contexts the locus of interaction might consist of a well-defined set of households who live within a given geographical space such as a street (see, for example, Guinnane, Moehling, and O'Grada (2001)).

6 Econometric Models of Fertility Behaviour

In order to motivate a modelling strategy we first consider the observed data. For each woman we observe $\{C_i, \mathbf{x}_i\}$, where $C_i \in \mathfrak{c} \subseteq \{0, 1, 2, ...\}$ represents a count of the total number of children ever born, and $\mathbf{x}_i$ is a vector of characteristics which includes individual characteristics, together with ethnic, cluster, and household attributes. We let $\mathbf{x}_i^I$ denote a vector which includes only individual characteristics.

At the outset we recognise that a standard linear model is inappropriate given that count data is both integer valued and heteroscedastic by construction.[15] In section 6.1 we examine a baseline Poisson model for the counts of children ever born, and determine the appropriateness of the attendant distributional assumptions. In section 6.2 we focus on the specification of the conditional mean, and in the light of our theoretical model presented in section 4, we propose an extension of the standard atomistic specification by accounting for social interaction at the level of the household, the village cluster and the ethnic group.

6.1 A Baseline Poisson Model

The Poisson distribution for the number of children ever born $C = 0, 1, 2, ...$, and for intensity parameter μ, is given by

$$f(C_i|\mathbf{x}_i) = \frac{e^{-\mu_i}\mu_i^{C_i}}{C_i!}, \quad C_i = 0, 1, 2, ... \tag{9}$$

Given the nonnegativity constraint, the Poisson density in (9) is often accompanied by an exponential conditional mean function $E(C_i|\mathbf{x}_i) = \mu_i = \exp(\mathbf{x}_i'\boldsymbol{\beta})$. Although in modelling the determinants of the number of children ever born to a woman, the Poisson density represents a natural benchmark, a number of specific features of both the population and the sample suggests that it will be necessary to modify the basic model to accommodate particular departures from conditions which satisfy a Poisson process. We begin with the violation of the assumption of conditional independence.

A violation of the assumption of conditional independence in the Poisson model with cross-section data may originate as a result of various forms of interactions across observations. However, unlike count process which are indexed by time, with an attendant immutable ordering, there is greater uncertainty as to the appropriate ordering. The key issue here is whether interaction is viewed as a nuisance process, or is integral to the substantive focus of the research.[16] Although the failure to account for dependencies across observations can result in inefficient estimation,[17] in this study our objective is to construct a structural model of the interaction processes by extending the Beckerian utility model for CEB to account for the impact of interactions across households which belong to distinct ethnic groups, and reside in

[15] This follows given that for any process bounded at zero, the variance will be an increasing function of the mean.

[16] King (1989) refers more generally to contagion as a process where the expected number of events at some time is dependent upon the realized number.

[17] See Moulton (1986), Chamberlain (1980) and Mundlak (1978).

clusters of households.[18] Subsequently we use this framework to determine whether such interactions generate different levels of CEB counts over the entire distribution of CEB, relative to that predicted if the only conditioning information was $\mathbf{x}_i^I$.

6.1.1 Neglected Heterogeneity: General and Specific

In our sample of married women between the ages of 15 and 49, the mean number of children ever born is 3.2, whilst the variance is 10.2, resulting in considerable overdispersion. As such one of the defining properties of the baseline Poisson model, namely the equidispersion property requiring that $E(C_i|\mathbf{x}_i) = Var(C_i|\mathbf{x}_i) = \mu_i$ is violated. In seeking to address this particular violation, we consider a specification which accommodates two possible departures from this baseline: a general form of unobserved heterogeneity, *and* a specific form which is manifest in the preponderance of excess zeros.

We extend the model in order to accommodate what we refer to as *general* unobserved heterogeneity by allowing the conditional mean to be determined by observed characteristics $\mathbf{x}_i$, and variation in an i.i.d unobserved heterogeneity component, ω_i. The conditional mean function is then written as

$$\begin{aligned} \mu_i(\omega_i) &= \exp(\gamma + \mathbf{x}_i'\boldsymbol{\beta} + \omega_i) \\ &= e^{\mathbf{x}_i'\boldsymbol{\beta}} e^{(\gamma+\omega_i)}, \end{aligned} \tag{10}$$

noting that since ω_i is unobserved, (10) suggests a random-effects interpretation of this particular extension. Conditional on ω_i, the CEB count C_i is distributed Poisson

$$f(C_i|\mathbf{x}_i, \omega_i) = \frac{e^{-\mu_i(\omega_i)}(\mu_i(\omega_i))^{C_i}}{C_i!}, \quad C_i = 0, 1, 2, ..., \tag{11}$$

with the unconditional probability given by

$$f(C_i|\mathbf{x}_i) = \int_0^\infty [e^{-\mu_i(\omega_i)}(\mu_i(\omega_i))^{C_i}/C_i!]g(\omega_i)d\omega_i. \tag{12}$$

To ensure a closed form expression for (12), a gamma distribution with mean 1 and variance α is chosen for $g(\varpi_i)$. The resulting mixture of Poisson and gamma components can be interpreted as the Negative Binomial model; the first two moments are $E(C_i|\mathbf{x}_i) = \mu_i$ (thereby ensuring the consistency of parameter estimates delivered by the PRM model in the presence of overdispersion) and $Var(C_i|\mathbf{x}_i) = \mu_i(1+\alpha\mu_i)$. One test for over-dispersion is a Wald test of whether α is different from zero.

One notable feature of count data, and fertility counts in particular, can be seen from Table 4, where we present a simple frequency distribution for children ever born. Although our sample includes 26% of women with no children, one characteristic of the baseline Poisson model is the tendency to underpredict a zero count (see for example, Poston and Mckibben (2003)). An excess zeros problem

[18] As such the fundamental characteristic which differentiates our approach from multilevel (variance components) techniques is that we assign, theoretically as well as empirically, a more prominent role to the mechanisms by which social interactions affect fertility behaviour.

may, in part, originate as a result of women in the sample who cannot have children, and for whom there are no reliable indicators of the infertile state. In this sense we might usefully think of the excess zeros problem in the context of CEB as originating from a *specific* form of unobserved heterogeneity in the form of a missing fertility indicator, say f_i. If such an indicator were available, and subsequently included as a component of the conditioning information $\mathbf{x}_i$, the excess zeros problem would not arise. However, given that f_i is unobserved a class of zero-inflated (or mixture) models have been proposed, where the vernacular *inflation* is used to denote the fact that the excessive probability mass at zero (relative to that allowed under a standard PRM) is generated by infertile women who are otherwise similar to fertile women choosing to have children.

We note that there are a number of alternate approaches to deal with this particular mixture model. Guinnane, Moehling, and O'Grada (2001) point out the hurdle Poisson model is predicated on the assumption that excess zeros are generated by the presence of a fixed cost associated with the count process. However, the major distinction between the two approaches is that the hurdle specification assumes that once the first hurdle has been crossed, all realizations from a Poisson density are strictly positive.[19] In contrast the mixture model assumes that the data are drawn from two different populations: an always zero CEB count population (with probability p) and a population (with probability $1-p$) that may include both zero and positive CEB counts.

6.2 Conditional Mean Specification

We turn our attention to the specification of the conditional mean function, and specifically integrate the econometric specification with the theoretical model outlined in section 4. We write the extended conditional mean for the baseline Poisson model as

$$E(C|\mathbf{x}^I, \mathbf{h}, \mathbf{p}, \mathfrak{e}) = \exp(\mathfrak{e} + \beta^e E[f(C|\mathfrak{e})] + \beta^c E[f(C|\mathfrak{cl})] + \boldsymbol{\gamma}\overline{ed}_{cl} + \mathbf{h}'\boldsymbol{\theta} + \mathbf{p}'\boldsymbol{\kappa} + \mathbf{x}^{I\prime}\boldsymbol{\eta}), \quad (13)$$

where $\mathfrak{e}$ denotes a categorical variable denoting ethnic affiliation, $E[f(C|\mathfrak{e})]$ represents some function $f()$ of the expected number of children ever born at the level of the ethnic group. $E[f(C|\mathfrak{cl}]$ captures similar effects which are localised in the sense of being confined to women who reside in the same cluster of households. $\overline{ed}_{cl} = \frac{1}{n_{cl}} \sum_{i \in cl} ed_i$ with ed_i denoting the education level of the i^{th} woman residing in the same cluster. $\mathbf{x}^I$, $\mathbf{h}$, and $\mathbf{p}$ denote, respectively vectors of individual, household and cluster attributes; and $\boldsymbol{\eta}$, $\boldsymbol{\theta}$, and $\boldsymbol{\kappa}$ are the associated vectors of unknown control parameters. β^e, β^c, and $\boldsymbol{\gamma}$ are scalar parameters denoting, respectively, endogenous and exogenous effects.

6.2.1 Endogenous and Exogenous Interaction Effects

[19] See Mullahy (1986).

We represent endogenous effects attributable to the fertility choices of other women in the same ethnic group and cluster using

$$\beta^e E[f(C|\mathfrak{e})] = \partial S_2^{\mathsf{e}}(n_i, \overline{n}_{-i}^{\mathsf{e}})/\partial n_i = \beta^e(\overline{n}_{\mathsf{e}_{-i}} - n_i), \tag{14}$$

$$\beta^c E[f(C|\mathfrak{cl})] = \partial S_2^{\mathsf{e}}(n_i, \overline{n}_{-i}^{\mathsf{cl}})/\partial n_i = \beta^c(\overline{n}_{\mathsf{cl}_{-i}} - n_i), \tag{15}$$

where n_i denotes the number of children ever born to woman i, and $\overline{n}_{\mathsf{e}_{-i}}$ ($\overline{n}_{\mathsf{cl}_{-i}}$) is the mean CEB for ethnic group (cluster). β^e represents the cross-partial effect[20] $\frac{\partial^2 S_2^{\mathsf{e}}(.)}{\partial n_i \partial \overline{n}_{-i}^{\mathsf{e}}}$. β^c is similarly defined.

We represent exogenous effects attributable to *observing* the education of other women in the same cluster using $\overline{ed}_{cl}$. This is in keeping with our discussion earlier of the effect of education as iconic, with some individuals aspiring to the attributes of other higher educated members within their cluster. We choose to model these education interactions at the level of the cluster as higher educated women in the locality (but not necessarily in the household) may be leaders and opinion-makers who influence others who are less educated to adopt low fertility norms. Hence we consider the proportion of women in the cluster who had completed higher education as representative of this exogenous education effect on fertility. It is important to clarify here that we are using a proportionate metric for the exogenous effect as we believe that the average level of education has an influence on fertility behaviour due to the 'iconic' value of education discussed earlier, rather than the distance of a particular woman's educational attainment from the attainment of better-educated women in the group. Therefore we argue that the proportionate metric is a better representation of the woman's actual behaviour than the conformist metric, at least with reference to the effect of education on fertility outcomes.

Above we defined a number of endogenous and exogenous effects specified at the level of the village cluster. It is important to include other measures of cluster *attributes* if we are to reduce the likelihood of spurious inference. Critical in this respect is access to water and fuel infrastructure. Based upon the work of Dasgupta (1993), better access to water and fuel reduces the demand for child labour to collect them, and that this is turn reduces the demand for children, and hence the fertility rate (Iyer (2002); Aggarwala, Netanyahu, and Romano (2001)). Access to fuel was measured by a mean-level effect for access to electricity. Access to water infrastructure at the cluster-level was measured according to whether access to water was located (i) in the residence (either piped into the residence or a well was located in the residence); (ii) obtained from a publicly-provided source (either a public tap or a public well); (iii) obtained by collecting it from a river, stream, pond or lake; or (iv) whether rainwater was relied upon as the chief source of water. In addition, another cluster-level attribute included was whether or not a radio was listened to at least once a week which again was aggregated up from whether or not an individual listened to a radio. The reason that we use cluster-level variables is that we believe that there may be an aggregate effect which is an attribute of the cluster which exerts an impact over and above the effect at the level of the individual.

[20] The notation $S_1()$ and $S_2()$ refers, respectively, to the proportionate and conformist social utility functions presented in Section 4.

6.2.2 Household-level Interactions

At the level of the household there are a number of channels of interaction that are relevant for fertility decisions. One measure of interaction is the practice of polygyny: 30% of households in the sample had more than one wife living in the household. Polygyny can have two impacts on fertility: first, it can increase fertility if there is more help provided with child-care, creating a micro-level externality within the extended family household. Second, if there is greater discussion of family planning issues, then this might work to reduce fertility. Alternatively if high-fertility norms are espoused within the extended household then the effect might be to increase fertility.[21]

The number of usual residents in the household, where usual residents excludes the number of children-ever-born to women, was also included as another possible channel of interaction. The rationale for the inclusion of this variate is that if other household members help with child care, then this considerably alters the benefits and costs of child bearing to parents. Where kin help with care, or indeed as with other African societies where 'fosterage' is common, other residents have an important bearing on a woman's total fertility. This variable is also important because such interactions may be manifest in the form of discussions about family planning. If, for example, a woman was likely to discuss family planning with other residents, her decision-making is more likely to be influenced by them. A number of other demographic studies of poor societies (see, for example, Iyer (2002)) have found that the role played by residents within the household or within close proximity, such as friends and neighbours, is important for fertility.

One of the issues that we were also interested in at the level of the household and which has relevance for the effects at different levels is the age of the women. The household-level effects were interacted with variables for women who were 'young' (ages 15-34) compared to women who were 'old' (ages 35-49). This was done because it was expected that some of the variables such as measures of household status might be more relevant to the decision-making of younger than older women, addressing an inherent time-inconsistency problem in the DHS data Arulampalam and Bhalotra (2003).

A number of household level variables were included as controls for income and other characteristics. One notable problem here is that in the DHS no direct questions on income were asked of survey respondents. Utilising a series of questions asked about household quality, access to infrastructure, and the ownership of consumer durables, it was possible to construct a number of indicators which control for the economic status of households. The quality of roof construction was measured on the following (increasing quality) scale: iron (mabati), tiles, grass or thatch, and other material. The quality of the floor was measured using (decreasing quality) scale: mud, dung or sand, wood planks, tiles or polished wood, and cement. Other indicators of the quality of housing infrastructure were whether or not it had a toilet, and the number of rooms for sleeping.

[21] A number of recent studies have examined the role of polygynous marriage norms in influencing demographic behaviour. See for example Hogan and Biratu (2004)).

The status of the household may also be proxied by ownership of consumer durables. Such items provide a relatively good measure of income given that the purchase of consumer durables is not subject to short-run variability, and as such is more representative of the longer-term status of the household. The variables that were used are whether or not the household owns a radio, a television, a telephone, and a bicycle, the latter being important for access to fuel and water. Note however, that ownership of a radio and a television could also be seen as indicative of the household's access to the mass media.

6.2.3 Individual-level Effects

The choice of variables to control for individual level effects is in keeping with a long lineage of economic models of fertility decisions. These variables are age of the woman measured in years and an age-squared variable accounting for the non-linearity associated with age-related variables. This controls for older women having, on average, more children than younger ones. The education of the woman is included and measured as the highest level of education attained, separating out primary, secondary, and higher education effects. The influence of the media was measured as whether the woman listened to the radio at least once a week. This variable was included on the assumption that greater information about contraceptive technology would be available to the woman if she listened to the radio.[22] A binary variable recorded how long the woman had lived in the community.[23] This variable was included on the assumption that if a woman has resided in the community for a longer period, then she was more likely to have formed stronger networks in the cluster, which in turn may influence her fertility behaviour. Similar to the household-level effects, the individual-level effects for education and for listening to the radio were also interacted with the dummies for younger and older women to address the time inconsistency problem with DHS data.

In both (9) and (11) the Poisson density is represented based upon the implicit assumption that the probability of each woman of having one or more children is the same. Given that our sample includes women between 15 and 49 years of age, we account for the fact that for younger women we effectively observe truncated fertility lifetimes by adding the variables 'age' and 'age squared' to control for the fact that older women would have completed their childbearing, and therefore might display higher fertility, relative to younger women in the sample. Another method that is used in the present study to deal with the time inconsistency problem in using DHS data is to interact dummy variables for older and younger women with other model effects to take account of the fact that this may reflect the distinction between completed and incomplete fertility. The specific nature of the time inconsistency problem is that women may be making decisions with respect to their fertility considerably before the time when they were selected for survey and interview, and

[22] Messages about this are routinely broadcast on the national radio channel, the 'Voice of Kenya' in English, Swahili and the vernaculars.

[23] If she had lived in the same community since 1993 then this variable took the value 0; if she had migrated in the interim period and had therefore lived in at least two communities, then this variable took the value 1.

that therefore their current economic status (as for example measured by ownership of consumer durables, or the current state of the roof quality or floor quality in their homes, may not be as representative of their economic status *at the time that they made their decision about having a child* (or having additional children). This creates the time inconsistency problem in DHS data which has been pointed out by Arulampalam and Bhalotra (2003).

7 The Identification Problem

In considering the problem of the identification of endogenous and exogenous effects, we consider data $\{C, \mathfrak{g}, \mathbf{x}\}$ where C is a scalar outcome measuring children ever born (CEB), $\mathfrak{g}$ is a categorical variable indicating group membership, $\mathbf{x}$ is a $k \times 1$ vector of characteristics recorded at the level of the individual, including educational attainment, age, and income. u denotes a vector of unobserved individual characteristics.[24] A linear model of CEB which considers different manifestations of social interaction is written as

$$C = \alpha + \beta E[C|\mathfrak{g}] + E(\mathbf{x}|\mathfrak{g})'\boldsymbol{\gamma} + \mathbf{x}'\boldsymbol{\eta} + u, \tag{16}$$

where α and β are unknown scalar parameters,[25] and $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$ are, respectively, $k \times 1$ and $l \times 1$ parameter vectors, with $l \leq k$. Conditioning on $\mathfrak{g}$ restricts the expectation over members of the appropriate group. The genesis of the identification problem derives from the simple observation that in (16) we specify a model for children ever born which is linear in a mean endogenous effect, $E(C|\mathfrak{g})$, and a mean exogenous effect, $E(\mathbf{x}|\mathfrak{g})$. It therefore follows that the reduced form representation of the conditional mean $E(C|\mathfrak{g}, \mathbf{x})$ will be a function of $E(\mathbf{x}|\mathfrak{g})$.[26] This is obviously central to the question of identification, namely we start with the intention of disentangling a mean group (ethnic) effect, in terms of the three components $E(C|\mathfrak{g})$, $E(\mathbf{x}|\mathfrak{g})$, and $E(u|\mathfrak{g}, \mathbf{x})$, and observe that, not surprisingly, the reduced form $E(C|\mathfrak{g}, \mathbf{x})$ is a function of $E(\mathbf{x}|\mathfrak{g})$.

There are a number of ways to circumvent the identification problem. Within a linear framework the simplest yet probably the most restrictive method, is to impose a full set of zero restrictions on $\boldsymbol{\gamma} = 0$, thereby adopting the assumption that any interactions are mediated solely by endogenous effects. We note that such an approach is potentially problematic for making inference on multiplier effects, since such restrictions may result in biased estimates of pure endogenous effects. This is important since policy proscriptions which follow from endogenous versus exogenous effects are likely to be very different. However, the upside of such an approach is that these restrictions automatically generate a possible set of instruments ($E(\mathbf{x}|\mathfrak{g})$).[27]

[24] Note that at this juncture we focus simply on the question of whether it is possible to separately identify a single pair of exogenous and endogneous effects. The question of whether more than one type of endogenous effect is identified is considered in section 9.

[25] For clarity of exposition we restrict β to be a scalar quantity. However, in the empirical section we allow endogenous effects to vary across ethnic groups such that β is now a vector.

[26] See Manski (1993) for further details.

[27] See, for example, Case and Katz (1991) and Gaviria and Raphael (2001). Brock and Durlauf

In this study identification is achieved as a consequence of a correctly specified functional form. Namely, in models of fertility choice we apply a link function to the conditional expectation $E(C|\mathfrak{g}, \mathbf{x})$ so as to correctly represent the nature of the observed data. Identification is then achieved as a result of the specified nonlinear model, rather than as a specifically targeted means to blur the linear dependence between endogenous and exogenous effects. The conditional mean is then written as

$$E(C|\mathfrak{g}, \mathbf{x}) = F(\alpha + \beta E(C|\mathfrak{g}) + E(\mathbf{x}|\mathfrak{g})'\boldsymbol{\gamma} + \mathfrak{g}'\delta + \mathbf{x}'\eta), \tag{17}$$

where $F(.)$ is a nonlinear monotonic transformation. As a consequence we note that although the conditional mean $E(C|\mathfrak{g})$ will still be a function of $E(\mathbf{x}|\mathfrak{g})$, the linear dependence between these two objects is automatically removed given that the we have imposed a restriction on the support of $E(C|\mathfrak{g})$: between zero and one for a simple binary choice model of fertility, and between zero and $\varkappa$, where $\varkappa$ is the maximum number of children ever born in either a multinomial[28] or count model representation. Subsequently, as long as the support of $E(\mathbf{x}|\mathfrak{g})$ is sufficient to reveal the nonlinear dependence between an endogenous and an exogenous effect, Manski's reflection becomes blurred and identification is achieved.

Identification problems may also arise as a consequence of endogenous group formation. In this instance a correlated effect may also be manifest if group formation is based on unobservable characteristics. Here our identification assumption is based on the belief that individuals do not self-select into the groups we examine. They do not self-select at the level of the household. At the level of the cluster there is selection insofar as individuals locate in a cluster due to their ethnicity, however this is exogenous to fertility behaviour. We also control for ethnicity.

8 Endogeneity Bias

In allowing for the possibility that interaction effects on fertility choices operate both through exogenous and endogenous interactions, we noted that the specification of a nonlinear count model breaks the linear dependence between the mean exogenous and mean endogenous effects. However, although both these effects are identified in the population, the question of how to deal with the obvious bias in estimating the endogenous effect remains. Endogenous cluster conformist effects are measured as $(\overline{n}_{\mathsf{cl}_{-i}} - n_i)$, where n_i, children ever born to woman i, is the dependent variable in a non-linear Poisson model. In the most general sense we postulate an effect stemming from the average fertility behaviour of a group to the individual. But logically, individual behaviour also affects the action of the group, generating a simultaneity problem independent of whether i is in the cluster or not.

In this study we circumvent the endogeneity using information on fertility preferences. The use of stated preferences for fertility here is linked directly to the

(2000) take a less extreme approach and achieve identification by locating an individual level variable whose group level average can be ruled out a priori as an exogenous effect. See Kawaguchi (2004) for an approach which utilises subjective perceptions of peers' average behaviour, instead of average behaviour.

[28]See Brock and Durlauf (2003) for a discussion of multinomial choice with social interactions.

observation by Manski (1993) that the estimation of endogenous interaction effects is tenuous if based solely on observed behaviour. The use of stated preferences as an instrument[29] is based on two arguments. First, stated and revealed preferences are correlated to the extent that actual fertility reflects preferences about ideal fertility, which is the number of children women want in the absence of any kind of method to control family size, or other constraints. Second, although fertility decisions within a group (and therefore mean group fertility) are affected by realised individual fertility decisions, stated preferences are exogenous to mean group fertility behaviour. For example, the survey data provided information on respondents revealed and stated preferences for family size. With respect to stated preferences, if the woman had at least one living child, then she was asked, '*If you could go back to the time you did not have any children and could choose exactly the number of children to have in your whole life, how many would that be?'* Conversely if the woman had no living children she was asked, '*If you could choose exactly the number of children to have in your whole life, how many would that be?'*

In this study observed differences between stated preferences and realised fertility choices reflect two components. First, there is the standard problem which plagues intentions data, namely the assumption that elicited choice expectations coincide with actual choices that respondents would make if the posed scenario were realised. In addressing this point, we note that we do not require that individual level differences between stated and revealed preferences average out.[30] Rather given that we are making use of the stated preferences as an *instrument*, we simply require that this measure provide independent exogenous variation for the endogenous revealed fertility preferences.

Second, these differences are also determined by the extent to which complete fertility lifetimes are observed, and the issue that very few women might say that their ideal family size is 'zero'. This is one reason why in the econometric analyses conducted for this study and reported in later sections of this paper, that we specifically use women's age as a crucial variable to interact other regressors with, as we might expect fertility scenarios for older women to be different compared to those for younger women. We also note that the wording of the question implies different hypothetical scenarios for different subgroups. For example, although the estimate of mean unwanted fertility, measured as the difference between revealed (C) and stated preferences (I) for fertility is relatively small ($\bar{I} = 3.98$ and $\bar{C} = 3.19$) differences across subgroups are noteworthy. For older women realised ceb ($\bar{C_O} = 7.05$) is on average significantly higher than ideal family size ($\bar{I_O} = 4.95$), which is consistent with the findings of other studies in terms of a high degree of 'unwanted fertility'. For younger women with incomplete fertility we find that, as expected, ideal family size exceeds realised fertility ($\bar{I_Y} = 3.82$ and $\bar{C_Y} = 2.47$).

[29] For other examples of the use of stated preference data in studies of social interaction see, for example, Kawaguchi (2004) and Delavande (2004).

[30] Make reference to Manski's work on eliciting future intentions based upon probabilistic expectations to explicitly handle the uncertainty on behalf in the decision maker in providing a response.

9 Data and characteristics of DHS survey

Reflecting the population breakdown in the country as a whole, our sample, taken from the Kenyan Demographic and Health Survey (KDHS)Ken (1999),[31] has data on the Kalenjin, the Kikuyu, Luhya, Luo, Kamba, Kisii, Mijikenda/Swahili and Meru/Embu.[32] The Kenya DHS adopted a two-stage stratified sampling approach that selected households located within primary sampling units (PSU) or sampling clusters. These clusters are identical to the complete enumeration of sample clusters which took place as part of the 1977 National Demographic Survey. The sample points themselves are identical to those chosen in the sampling frame maintained by the Kenyan Central Bureau of Statistics. In the 35 of Kenya's 42 districts that were included in the survey, there were 536 clusters - 444 rural clusters and 92 urban clusters, of which 530 were non-empty clusters.[33] The location of the clusters geographically are identifiable within the district and province (DHS and Macro International, 1999: 179-182). A complete list of all households in each cluster was recorded between November 1997 and February 1998. From the remaining 530 clusters, a systematic sample was drawn of, on average, 22 households in urban clusters and 17 households in rural clusters. This formed a total of 9465 households. In these households, all women age 15-49 were targeted for interview. Response rates varied by province from approximately 88% to 99% (for more details see KDHS, 1998 p.180).

In addition to the geographical stratification of data by cluster and household, it is also possible to stratify these data by ethnicity. As also discussed in section 2, the major Kenyan ethnic groups (covering 88% of the population) are Kikuyu, Luo, Luhya, Kamba, Kalenjin, Kisii and Mijikenda/Swahili. The Kikuyu is the largest, with 17.9% of the population, while Mijikenda/Swahili is the smallest, with 5.0%. The representation of each group in the sample is similar to its representation in the whole population. Table 4 depicts the total number of cases under study grouped according to region and ethnicity. The largest number of women included in the sample live in the Rift Valley region, while the smallest number of women sampled live in Nairobi. The nature of the sample is therefore particularly useful for analysing behaviour in rural Kenya. Looking at the distribution by ethnic group, the largest sample was drawn from the Kalenjin group, while the smallest was drawn from the Meru/Embu group.

It is important to emphasise that in many cases the cluster populations are homogenous with respect to ethnic group. As a result there will many cases very little residual variation on the ethnic endogenous effect after we have controlled for the endogenous cluster effects such that there will be insufficient independent

[31] The survey contains interview data for 7800 women aged 15-49. It contains women from all of Kenya's large ethnic groups and covers a wide geographical area, omitting only the areas of extremely low population density in the North. Geographically. Kenya is divided into 7 provinces which are further subdivided into 47 districts. The Kenya DHS covered 42 of Kenya's districts; 35 were sampled, 7 were not. 17 districts were oversampled.

[32] Our study excludes the Maasai, a pastoral group living in the Kenyan Rift Valley, and other tribes such as the Galla and Somalis, who live in the north-east.

[33] 6 of the clusters could not be included in the survey due to inaccessibility (DHS and Macro International 1999: 180).

variation in these two effects to separately (empirically) identify them. Section 11 discusses this in further detail.

According to the DHS data, a rural woman has on average about 5.2 children compared to fertility among urban women at 3.1 children. Fertility differentials by the level of education show that illiterate women bear on average 5.8 children compared to 3.5 children for women with secondary school education (DHS 1998, p. xvii). The sample reveals demographic differences between ethnic groups: for women aged 40-49, the mean CEB varies from 5.91 for Kikuyu to 7.56 for Luo. The differences in fertility by ethnic group are clearly very large, and are evident whether we are examining actual completed fertility (CEB), ideal number of children desired, or unwanted fertility. Table 2 shows the mean CEB grouped by region and by ethnicity for various age cohorts in the sample. CEB varies from a low of 1.7 in Nairobi/Central to a high of 3.3 in Nyanza. This pattern is also seen by region for differing age cohorts. The Kikuyu, Kalenjin, Luhya and the Luo follow similar patterns for all age groups. For example, CEB is low among the Kikuyu (in the Nairobi region), and high amongst the Luo (in the Nyanza region).

There are three-quarters of women and men in the sample either who want to limit births or to space them. Table 5 shows the stated ideal number of children desired by women in the sample. This is lowest in Nairobi/Central compared to Nyanza, western and coast regions. It varies from a low of 2.95 to a high of 4.36 in Coast. This pattern is also seen by region for differing age cohorts, except for the Coast region (discussed below). All ethnic groups depict relatively similar behaviour except for Miji/Swahili (in the Coast) who depict very high ideal numbers, and the Kikuyu who are very low.

If we are concerned about the influence of social interactions on fertility behaviour, empirically it is useful to examine women's unwanted fertility in relation to their actual fertility. The last column of Table 5 depicts the demographer's measure of unwanted fertility in the sample measured as the total number of CEB less the ideal number of children desired for women aged 40-49 years who have largely completed their fertility. This is lowest in the Central region (-0.5) and highest in the Coast region (-1.7). In all regions, women are having greater numbers of children than they desire, except among women in the Coast region where it is the reverse. All ethnic groups are having greater numbers of children than they desire, except among Miji/Swahili, where it is the reverse. For the age group 40-49, unwanted fertility varies from 0.1 in the Coast region to 2.4 in the Eastern region. These data suggest that social interactions may matter for fertility behaviour in this society, as women are having more children than they desire, reflected in these estimates of unwanted fertility.

The DHS data show that knowledge of family planning in Kenya is very high: 98% of women and 99% of men were able to name at least one modern method of contraception. There are 39% of women who use contraception and the most widely used methods are contraceptive injectables, the pill, female sterilization and periodic abstinence. Contraceptive use does however vary greatly by region: while there are 61% of women in the Central province who use contraception, only 22% of women in the Coast province do so likewise. Only 23% of women with no education use contraception compared to 57% of women with secondary education. Both government and private medical sources provide access to contraceptives.

In the data used in the present study we consider individual-level data only for the 5994 women residing in a rural cluster, and who did respond. Table 6 presents the counts of children ever born in the sample. This shows that there is heaping of observations, particularly on the zero count.

10 Results

Given that the Poisson regression model (PRM) generates consistent parameter estimates in the presence of either under or over-dispersion (see, for example, Cameron and Trivedi (1986)), we utilize the PRM as a convenient benchmark model. The results of the different model estimations are discussed as follows. Table 7 presents the Baseline Poisson Model with cluster controls, household controls and individual-level effects included. Table 8 presents the Zero-Inflated Poisson model with CEB instrumented, and the social interaction effects together with the same set of controls as in the benchmark PRM. The dependent variable in the model is the total number of children ever born (CEB). Parameter estimates may be interpreted as the proportional change in the number of children ever born due to a unit change in the regressors.[34] Since the predicted count is given by $e^{\mathbf{x}'\hat{\boldsymbol{\beta}}}$, we also report the factor change in the expected count for a unit increase in a given x_k, denoted $e^{\hat{\boldsymbol{\beta}}}$, and emphasise this statistic when interpreting our results. Standard errors are calculated using the Huber (1967) method to account for intra-cluster correlation.

In discussing the results it is important to consider the significance of particular effects over the respective levels: individual, household, cluster, and ethnic. Consider the following hypothetical example. At the level of the cluster access to the media is measured by the proportion of households which had access to a radio; at the level of the individual, media access is measured by whether or not the individual listens to a radio at least once a week. Assuming that both of these effects are significant, our inference would be that after controlling for the negative effect of individual use of a radio, a woman living in a cluster which had a higher proportion of radios would experience a further reduction in her CEB. The interpretation for such a cluster effect is based on the observation that information about family planning is disseminated by radio, this information influences ideas about fertility decline, and that as the proportion of women in the local vicinity that listen to the radio increases, this creates a subsequent multiplier effect. In terms of population policy, this is important as it would suggest that not only are the characteristics of the individual women important, but also the characteristics of the clusters in which these women live. As we discuss below, in order to guard against spurious inference at the level of cluster interactions, we obviously need to ensure that we control for

[34] This follows since

$$\frac{\partial E(C|\mathbf{x})}{\partial x_j} = \beta_j \exp(\mathbf{x}'\boldsymbol{\beta})$$

where $E(C|\mathbf{x}) = \exp(\mathbf{x}'\boldsymbol{\beta})$.

income, since without such a control the inference might be misleading.

As noted above, one of the problems with the DHS survey is that no direct questions on income were asked of survey respondents. To address this deficiency a number of household level variables were included as controls for income and other characteristics. Measures of roofing quality are significant, suggesting that if the woman lived in a household that was wealthier, as measured by a better quality roof such as iron or tiles, then her CEB was more likely to be lower, relatively to the base category - a thatched roof. If a woman lived in a house which had an iron roof, the expected number of children ever born would be approximately 5% lower than if she was resident in a house with a thatched roof; the difference in fertility outcomes between women living in thatched-roof homes compared to those living in homes which had tiled roofs was as high as 23%. This result represents additional evidence of the importance of income as an influence on fertility. The presence of a television in the household also mattered for lowering fertility: women who lived in households which owned televisions had on average 8% fewer children. It is possible to argue that this is a measure of both an income effect and a measure of better media access, with both effects working in the same direction to reduce expected CEB.

As expected many of the individual-level variables were highly significant: these included the woman's age with an older woman exhibiting higher fertility compared to her younger counterpart. This is something we might expect if norms concerning family size have evolved over time. Women today are more likely on average to have fewer numbers of children compared to previous generations of women before them. To isolate further the influence of age on fertility, we interact some of the other variables with dummy variables for older versus younger women[35]. For example, at the level of secondary and higher education, education exerted a negative impact on fertility. But the magnitude of the impact on older compared to younger women is different. For older women, while primary and secondary education were not significant, higher education was significant for reducing fertility, with a factor effect of 20%. For younger women, the effects are even more dramatic: the effects of secondary (higher) education reduces fertility by, on average 22% (36%). The base category was women with no education. Primary education has the expected positive relationship with CEB. If the woman had lived continuously in the community since 1993, then this was likely to reduce her ideal number of children by about 4%. One reason for this finding may be that women who stayed more continuously in one region were more likely to benefit from the impact of the greater 'social capital' generated by living in the community for long periods. It is also possible that they are more likely to form stronger networks with others if they have lived there longer, and this would be particularly important for fertility-related issues.

Since we are not seeking to isolate the effects of exogenous or endogenous interactions, the inclusion of ethnic dummies allows us to capture a composite ethnic effect on fertility. Using the Kikuyu as the base group, we find significant effects for the Kalenjin, Luhya, and Luo. In comparing our results with the unconditional fertility

[35] We take a threshold here as greater than 35 as 'older', on the assumption that a woman's reproductive life span is from 15 to 49, and as also assumed in other demographic analyses.

means in Table 1, we note that the magnitude of the attendant factor effects coincide with the ranking of ethnic groups with high fertility. For example, our inference here would be that an otherwise identical Kikuyu women would experience, on average, a 15% increase in fertility if she were exposed to the norms and behaviours of the Luo ethnicity. In moving now to discuss the results of an extension of the PRM model, one of our objectives is to decompose these fertility differences into a number of exogenous and endogenous ethnic effects. Finally we note that, using Western as the reference province,[36] all regional effects are insignificant with the exception of the Central province.

The Baseline Poisson model represents a valuable point of departure in demonstrating the importance of the individual-level factors. However, the model maintains an equidispersion assumption, is not able to adequately model a preponderance of the zero counts, and does not take into account the effect of endogenous and exogenous interaction effects at the level of household and cluster. Utilising the discussion in section 6.1, we take the PRM as the null model and first test for departures from equidispersion based upon a *general* form of unobserved heterogeneity as represented by the negative binomial model with variance $Var(C_i|\mathbf{x}_i) = \mu_i(1 + \alpha\mu_i)$. Utilising both a Wald and likelihood ratio test ($\mathrm{H}_0 :\ \alpha = 0$, $\mathrm{H}_1 :\ \alpha > 0$) of whether the PRM model should be extended to account for a *general* form of unobserved heterogeneity was conducted, we were not able to reject the null at any reasonable significance level. As a consequence we extend the baseline Poisson model to account for the *specific* form of unobserved heterogeneity - namely the fact that we do not observe an infertility indicator - is supported by the data.

10.1 The Zero Inflated Poisson IV model with social interaction effects

Table 8 presents the results of a Zero-inflated Poisson regression with CEB instrumented[37] by the ideal number of children estimated on the Kenyan data, with cluster-level endogenous and exogenous effects, household effects and individual-level effects.[38] Cluster–level variables are divided into cluster-level social interaction terms and cluster-level controls. The cluster-level interaction terms encompass endogenous and exogenous effects.

Endogenous effects The localised endogenous effects emanating from within the village cluster, are captured by a parameter which weights the variable $(\overline{C}_{cl} - C_i)$ - namely the difference between a mean-level and individual CEB (instrumented by ideal family size) for women who live in the same cluster. We observe that in many cases these parameter are significant and positive. In differentiating these

[36] Since we have focussed our study on rural clusters, Nairobi is excluded from the analysis.

[37] The instrumental variable regression is a Poisson model using ideal family size (*ifs*) as an instrument for children ever born. The parameter value and standard error for *ifs* were, respectively 0.0218 and 0.004. Note that a factor effect of 1.022 is consistent with the observation of unwanted fertility in table 5. A full set of individual, household, and cluster level exogenous variables were also included.

[38] Although the model contains a large number of variables based upon cluster aggregates, standard errors are again adjusted to account for any residual intra-cluster correlation.

effects according as to whether the woman was 'young' or 'old', we found that across all ethnic groups the endogenous effects were very important for younger women, on average, compared to the older women: factor effects ranging from 1.043 to 1.11 which suggests an increase in fertility of between 4% - 11% across ethnic groups. For older women, the effects wre significant for the Kisii, the Mijikenda-Swahili and the Luhya, with factor effects ranging from 1.014 to 1.052 - or approximately between 1% and 5%.

A key aspect of these results is that consistent with our anthropological understanding of the ethnic groups (and as discussed in section 3), for those clusters in which individuals were related both by blood and marriage, as for example, among the Kikuyu *mbari*, the endogenous effects on fertility are the strongest of any of the ethnic group effects on fertility.

Exogenous effects We consider exogenous effects as attributable to the effect of *observing* the level of education of other women in the same cluster. We measure this effect using the proportion of women who had completed higher education, allowing parameters to differ by ethnicity and the age of the woman[39]. Among the younger women, these effects were significant for the Kikuyu. The most significant negative effect was observed for younger Kikuyu women where the exogenous effect of education multiplies the expected number of children ever born by a factor of 0.105. Once again we observe that those groups who are related both by blood and by marriage, as for example the Kikuyu, exhibit significant and large interaction effects. The importance of these effects is that they suggest more clearly the channels through which the effect of ethnicity impacts on fertility. This is important for demographic analyses per se because compared to the results in Table 7 which represent the more conventional manner in which fertility anlaysis is conducted, the results in Table 8 show the importance of both including the significant ethnic effects on fertility that are the outcome of interactions between individuals, and the fact that they are important over and above the significance of individual and household characteristics. The effects in Table 8 are significant after controlling for the individual levels of educational attainment. For the older women however, these effects are not significant for any of the groups. This might suggest that the interaction of social interaction effects and education effects might be more important for younger cohorts of women, who increasingly learn from the better-educated members of their ethnic group more, as more and more women in these groups have achieved better education levels over time.

In order to reduce the likelihood of omitted cluster attributes confounding our inferences on exogenous and endogenous interactions a range of cluster-level controls are included: these are media access and access to water and fuel infrastructure in the cluster. Among these cluster-level controls, the effect of the media was significant and exerted a negative effect on fertility, with a factor effect of 0.892, or approximately 11% decrease. None of the mean-level cluster effects for access to

[39] For the Mijikanda-Swahili and Kamba ethnic group the numbers of women who had completed higher education education was extremely small across most clusters, and as a result, the exogenous education effect was constructed utilising the proportion who had completed a secondary school education.

electricity or water were significant.

Household-level effects At the household level we make a distinction between interactions within the household which may have implications for fertility, and household-level controls. Household-level controls are included so as to reduce the possibility that any inference on household-level interaction effects is not confounded by household-level omitted variables. The main household-level social interaction effect that was significant was if the woman was a usual resident, then this impact was more significant for older women, for whom it exerted a negative effect, with a factor effect of 0.893. The 'usual resident' variable captured the number of other residents in the household excluding the number of children at the time of the survey. The impact of polygyny was not significant either for younger or older women in the sample.

A number of household-level attributes are used as controls. In general we would expect that income would be negatively correlated with fertility, and this hypothesis was borne out by income indicators used in the model. For example, if the woman belonged to a household which had an iron roof or a tiled roof (relative to the base of straw or thatch), then the woman had, on average, lower fertility. In terms of orders of magnitude, having a tiled roof, for example, reduced the likelihood of a woman's fertility by approximately 29%; having an iron roof reduced the likelihood of a woman's additional fertility by approximately 6%. This finding is unsurprising as roofing quality is a strong indicator of income in a rural household. The floor quality variables were also significant: if the woman lived in a house which had a cement floor (relative to the base of sand), then this multiplied average fertility by a factor of 0.95. Two variables included in the model which could be construed both as income indicators and as controls for media access are radio and television ownership. Radio ownership was insignificant, but television ownership was highly significant. For example, ownership of a television reduced a woman's fertility by approximately 9%.

Individual-level effects The age of the woman depicted the expected relationship with fertility: older women, on average, have higher fertility than younger women. Primary education exerted a negative effect on fertility. This variable was important (at the 10% significance level) for the younger women, although not for the older women in the sample, with a factor effect of 6%. Secondary and higher education were significant for both older and younger women, and displayed the expected negative signs. If a younger woman was educated to the secondary level, she depicted on average 38% lower fertility; if she had had higher education, then this reduced her expected fertility by about 50%. For older women, these figures were 10% and 28% respectively. If an older woman had been resident for at least five years in this region, then again this had a negative impact on her fertility, and reduced her fertility, on average, by approximately 11%.

10.1.1 Inflation sub-model

Given that we do not observe an infertility indicator we are not able to determine whether any woman should be permanently assigned to the always-zero regime.

Subsequently, we utilise the inflation sub-model to estimate the probability that a woman has zero children due to being infertile. This probability was estimated using a simple probit equation. Given that we do not observe any reliable indicators of infertility we included a set of ethnic dummies, again with Kikuyu as the base ethnic group. The rationale for this specification is that since infertility is determined, in part, by a number of unobserved factors such as genetic composition, nutritional status, it is likely that ethnicity may represent some of this variation. Another factor that may be relevant here, which has been highlighted by medical anthropological studies (see Tanaka (2000)), is the practice of male and female circumcision, which is practised by most ethnic groups in Kenya. This practice may lead to lower fertility if it endangers the health of women. The findings of the inflation sub-model show significant ethnic effects for three groups: the Kisii, the Meru-Embu and the Luo. The findings for the Kisii and the Meru-Embu suggests that there are, relative to the Kikuyu, a significantly smaller proportion of infertile women in the sample for these two groups. For the Luo, the findings show that there are a significantly higher proportion of infertile women in this group compared to the Kikuyu, although this result was only significant at the 10% level. The findings from the inflation sub-model suggest that if we need to understand fertility behaviour in this society more fully, then we may need to look at differences in the degree of infertility amongst ethnic groups, coupled with the anthropological evidence, as an indicator of subsequent fertility behaviour. The preliminary findings from the inflation sub-model indicate that this may be a potentially important avenue for future research in this area, as high levels of fertility among some groups relative to others may be the outcome of a combination of factors including a predisposition towards infertility in some ethnic groups in this society. A joint test of the null hypothesis that all the parameters in the inflation sub-model were jointly equal to zero was firmly rejected[40].

11 Conclusion

Strategic complementarities in fertility decisions implies that a couple's fertility decisions may be dependent on the actions of others in the vicinity or in the society more widely. The mechanisms through which these complementarities occur are through social interactions. This paper has examined the importance of social interactions in the context of fertility behaviour in Kenya. We have examined the role of dependencies within a household; dependencies across individuals that reside in a cluster of households who locate according to ethnicity; and more globally ethnicity. By identifying the multiplicity of these channels, we have a better basis upon which to attempt to influence policy. More significantly, the existence of multiple channels of social interaction imply that an analyst attempting to isolate these pathways, needs to be cautious about the possibility of erroneous inference.

The key message that emerges from this study is that since fertility is influenced by interactions formed at different levels, then from a policy perspective, a specifically targeted intervention in an area of high fertility, or towards specific groups in

[40] The Vuong test statistic has an asymptotic normal distribution. The value of the test statistic was 3.60, with a p-value of 0.000. See Vuong (1989) for details of the form of this non-nested test.

society, arguably will be more effective than a 'one size fits all' approach to population policy. For example, one of the key findings of the study is that the exogenous effects of education vary significantly by ethnic group, and are seen to be important with large multipliers. In particular, the exogenous effects significantly reduced Kikuyu fertility. One reason for these findings may be the anthropological fact of the Kikuyu clusters in which individuals were related both by blood and by marriage, leading to significant exogenous education effects. But these types of findings do seem to suggest that a more nuanced approach to population policy is needed than simply a policy that aims at increasing education in general. It may be that the particular forms of education provided need to be context-dependent recognizing that education may have different impacts on people of different ethnic backgrounds. While this study did not explore these themes in depth, we believe that this is an avenue for potential future research in Kenya, and would be worthy of investigation especially from the perspective of population policy.

This conclusion is reinforced by the analysis of the endogenous effects on fertility. Women of different ethnicities are influenced variably by the observed fertility behaviour of others in their ethnic group. This varies most particularly by age. For example, the endogenous CEB effects were important for younger women of all ethnic groups in Kenya, but only important for some of the older women. This seems to suggest a gradual process of change in how much Kenyan women are being influenced by others in their group over time. Younger women, perhaps because of recent proactive population and health campaign initiatives in this country, are increasingly being influenced by others in their immediate group of reference. For older women, it is more difficult for us to draw firm conclusions as the findings for these women need to be tempered by a recognition of the time inconsistency problem inherent in using cross-sectional demographic data. In any event, this is the reason behind why one of the channels that we highlight in the econometric models is the influence of women's age. Both endogenous and exogenous effects are more significant for younger women in the sample than for older women. So specifically, we argue that by using *highly educated women, from specific ethnic groups, and resident in local communities*, as 'leaders' to spread information among women about the benefits of lower fertility, this may be a good strategy to ensure fertility decline in Kenya more widely. Thus, we argue that the effects of age, education and ethnicity, when they do coincide in particular individuals, seem to engender a capacity for reducing fertility in Kenya that is noteworthy.

One of the most interesting aspects of looking more closely at the education effects is that the size of the multipliers at every level are very large. For *all* younger women in the sample the effect of completing a secondary education is to reduce expected fertility by 38%. When we now allow for exogenous interaction effects attributable to the iconic effect of education for women in the same cluster, then for some ethnic groups we observe even larger multiplier effects. For the Kikuyu, we estimate an additional significant decline in fertility which we attribute to the effect of the proportion of women who completed secondary education within the cluster. This finding underlines the importance of anthropological characteristics of ethnic groups in reinforcing the impact of the social interactions. These findings are also consistent with the theoretical observation made by economists elsewhere that when the social interaction effects are very significant, the size of the multiplier is very large (Horst and Scheinkman (2003b)). This seems to be evident particularly for the effect of education on fertility in this sample population.

From a policy perspective, the analysis conducted here also suggests that in Kenya, we will need to target the demand for children via influencing people's desires about fertility. As discussed, on the one hand, education is one potential avenue by which to do so. But, on the other hand, the role of the media through promoting knowledge about family planning is critical in order to influence 'ideas' about lower fertility norms. Economic characteristics are important, and measures of income are significant determinants of fertility. If high fertility is a consequence of the demand for child labour to collect water and fuel infrastructure, then targeting access to piped water in the home and electricity is critical.

The observation that clusters of households are formed almost exclusively on the basis of ethnicity precluded us from estimating separate ethnic and cluster conformists effects. However, this finding is informative from a policy perspective in the sense that fertility behaviour via conformity at the cluster level mirrors the same at the ethnic group level. The findings are also relevant as to how policy is conducted. For example, if cluster populations are coincident with ethnicity as they are in Kenya, then potentially it is easier to target population policy at one group in a localised setting. If however the clusters are mixed, then potentially this creates a difficulty for targeted fertility interventions that might vary considerably by ethnic group. In this respect homogenous clusters might make targeting population policy more tractable.

In addition, given a single observation per individual we are not able to make inference as to the permanence of the interaction effects. However, although similar inference problems relate to whether in using a single cross-section an estimated individual effect represents a steady state relationship or some form of transition, we argue that in the case of social interaction effects due to fertility behaviour, the relevance of the orderings - the household, the cluster, and the ethnic group - is important. Although some ethnic norms are, in all likelihood, immutable, others may change over time. Nevertheless, these norms do emerge in empirical analysis as important, despite controlling for the other socio-economic characteristics of individual members of these ethnic groups. The significance of ethnicity may imply therefore that ethnic group leaders can be targeted to influence norms.

Further research in this area might consider relaxing the assumption that endogenous and exogenous effects are additive; an argument might be made that they are multiplicative. This extension would allow us to capture that component of individual utility attributable to a fertility choice that is dependent upon the fertility choices of others, and which for example, varies as a function of the mean level of education in the cluster. For example, some groups with, on average, lower levels of education, might be more 'conformist' than others with respect to fertility norms. So we might want to ask 'Will the Kamba who have lower education levels, also be more 'amenable' to norm-guided behaviour than the Kikuyu, who have higher education?' This obviously suggests a more complex form of interaction, but, as we see it, this would be a potential avenue for future research in this area.

Nevertheless, the key conclusion of our research is this: if we wish to make the issue of population control in poor countries like Kenya a thing of the past, then we need to understand the contemporary dynamics of social groups with respect to fertility, and respond to them accordingly. Conceptualised in this way, we believe that operationalising the use of the social interactions methodology for identifying effects on fertility behaviour that are important over and above the individual's characteristics, is going to be paramount for policy. Until then, population control in developing countries such as Kenya will not be history.

References

(1970): *Population Census 1969*Republic of Kenya, Vol. 1, Nairobi.

(1999): *Kenya: Demographic and Health Survey*National Council for Population and Development (NCPD), Central Bureau of Statistics (CBS), Office of the Vice President and Minsitry of Planning and National Development (Kenya) and Macro International Inc. (MI), Calverton, Maryland.

Aggarwala, R., S. Netanyahu, and C. Romano (2001): "Access to Natural Resources and the Fertility Decision of Women: The Case of South Africa.," *Environment and Development Economics*, 6, 209–236.

Ajayi, A., and J. Kekovole (1998): "Kenya's Population Policy from Apathy to Effectiveness," in *Do Population Policies Matter? Fertility and Politics in Egypt, India, Kenya and Mexico*, ed. by A. Jain. New York: Population Council.

Akerlof, G. (1997): "Social Distance and Social Decisions," *Econometrica*, 65, 1005–1027.

Arulampalam, W., and S. Bhalotra (2003): "Sibling Death Clustering in India: Genuine Scarring vs Unobserved Heterogeneity," Department of Economics, University of Bristol, UK, Bristol Economics Discussion Papers 2003.

Ascadi, T. F., G. J. Ascadi, and R. A. Bulatao (1990): "The Context and Causes of High Fertility," in *Population Growth and Reproduction in Sub-Saharan Africa: Technical Analysis of Fertility and its Consequences*, ed. by T. F. Ascadi, G. J. Ascadi, and R. A. Bulatao, pp. 131–132, Washington, D.C. The World Bank.

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

Becker, G. S., and K. M. Murphy (2000): *Social Economics: Market Behavior in a Social Environment.* The Belknap Press of Harvard University Press, Cambridge, M.A. USA.

Berg-Schlosser, D. (1984): "Tradition and Change in Kenya," *International Gegenwart Band 3*, Paderborn: Ferninand Schoningh.

Borjas, G. J. (1995): "Ethnicity, Neighbourhoods and Human-Capital Externalities," *American Economic Review*, 85(3), 365–390.

Brock, W., and S. Durlauf (2000): "Interactions Based Models," in *Handbook of Econometrics*, ed. by J. Heckman, and E. Learner, no. 5. North Holland Press, Amsterdam.

Brock, W. A., and S. N. Durlauf (2001): "Discrete Choice with Social Interactions," *Review of Economic Studies*, 68, 235–260.

Brock, W. A., and S. N. Durlauf (2003): "Multinomial Choice with Social Interactions," National Bureau of Economic Research, Inc, NBER Technical Working Paper: 288.

Caldwell, C., and P. Caldwell (1987): "The Cultural Context of High Fertility in Sub-Saharan Africa," *Population and Development Review*, 13(3), 409–437.

Cameron, A. C., and P. K. Trivedi (1986): "Econometric Models Bsaed on Count Data: Comparisons and Applications of Some Estimators and Tests," *Journal of Applied Econometrics*, 1, 29–53.

Case, Anne, C., and F. Katz, Lawrence (1991): "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths," National Bureau of Economic Research, Inc, NBER Working Papers: 3705.

Chamberlain, G. (1980): "Analysis of Covariance with Qualitative Data," *Review of Economic Studies*, 47, 225–228.

Coale, A. J., and S. C. Watkins (eds.) (1986): *The Decline of Fertility in Europe*Princeton. Princeton University Press.

Conley, T., and G. Topa (2002): "Socio-Economic Distance and Spatial Patterns of Unemployment," *Journal of Applied Econometrics*, 17, 353–395.

Cooper, J. (1988): "Coordinating Coordination Failures in Keynesian Models," *The Quarterly Journal of Economics*, 103(3), 441–463.

Dasgupta, P. (1993): "Poverty, Resources and Fertility: The Household as a Reproductive Partnership," in *Alternatives to Capitalism*, ed. by A. B. Atkinson, pp. 207–243. St. Martin's Press, New York.

Dasgupta, P. (2000): "Population and Resources: An Exploration of Reproductive and Environmental Externalities," *Population and development review*, 26, 643–689.

De Wilde, J. C. (1967): *Experiences with Agricultural Development in Tropical Africa*, vol. II: The Case Studies. John Hopkins Press, Baltimore.

Delavande, A. (2004): "Pill, Patch or Shot? Subjective Expectations and Birth Control Choice," Research paper. Author contact Universidade Nova de Lisboa and RAND Corporation.

Durlauf, S. N., and J. Walker (1999): "Social Interactions and Fertility Transitions," Paper prepared for workshop on 'Social processes underlying fertility change in developing countries.', 28-30 Jan 1998, organised by the Committee on Population, National Academy of Sciences.

Egero, B., and M. Hammraskjold (eds.) (1994): *Understanding Reproductive Change: Kenya, Tamil Nadu, Punjab, Costa Rica*Lund. Lund University Press.

Fapohunda, B. M., and P. V. Poukouta (1997): "Trends and Differentials in Desired Family Size in Kenya," *African Population Studies/Etude de la Population Africaine*, 12(1).

Federici, N., K. O. Mason, and S. Sogner (eds.) (1993): *Women's Position and Demographic Change*Oxford. Clarendon Press.

Ferguson, K., and T. SrungBoonmee (2003): "Culture and the Economy in Kenya," Paper submitted for the 1st Annual undergraduate research symposium, University of South Florida, Spring.

Gaviria, A., and S. Raphael (2001): "School-Based Peer Effects and Juvenile Behavior," *Review of Economics and Statistics*, 83(2), 257–268.

Gellner, E. (1981): *Muslim Society*. Cambridge University Press, Cambridge, UK.

Ginther, D., R. Haverman, and B. Wolfe (2000): "Neighborhood Attributes as Determinants of Children's Outcomes: How Robust are Relationships?," *The Journal of Human Resources*, XXXV(4).

Glaeser, E. L., and J. A. Scheinkman (1999): "Measuring Social Interactions," Working Paper, Harvard University.

Guinnane, T. W., C. Moehling, and C. O'Grada (2001): "Fertility in South Dublin a Century Ago: A First Look," Yale University Economic Growth Centre Discussion Paper No. 838.

Hajnal, J. (1982): "Two Kinds of Preindustrial Household System," *Population and Development Review*, 8(3), 449–94.

Hank, K. (2001): "Regional Social Contexts and Individual Fertility Decisions: A Multilevel Analysis of First and Second Births in Western Germany," MPIDR Working Paper No. 2001-015.

Hogan, D. P., and B. Biratu (2004): "Social Identity and Community Effects on Contiagitive Use and Intentions in Southern Ethiopia," *Studies in Family Planning*, 35(2), 79–90.

Horst, U., and J. A. Scheinkman (2003a): "Equilibria in Systems of Social Interactions," mimeo Humboldt University of Berlin, Department of Mathematics.

——— (2003b): "Equilibria in Systems of Social Interactions," mimeo, Humboldt University of Berlin, Department of Mathematics.

Iyer, S. (2002): *Demography and Religion in India*. Oxford University Press, Delhi.

Iyer, S. (2002): "Religion and the Decision to Use Contraception in India," *Journal for The Scientific Study of Religion*, 41(4), 711–722.

Kawaguchi, D. (2004): "Peer Effects on Substance Use Among American Teenagers," *Journal of Population Economics*, 17, 351–367.

Kenyatta, J. (1961): *Facing Mount Kenya*. Secker and Warburg, London, (first published 1938).

——— (1966): *My People of Kikuyu*. Oxford University Press, Nairobi.

King, G. (1989): "Variance Specification in Event Count Models: From Restrictive Assumptions to a Generalized Estimator," *American Journal of Political Science*, 33, 762–784.

Kohler, H. P. (2001): *Fertility and Social Interaction: An Economic Perspective*. Oxford University Press, Oxford.

Manski, C. F. (2000): "Economic Analysis of Social Interactions," *Journal of Economic Perspectives*, 14(3), 115–36.

Manski, C. F., and J. Mayshar (2002): "Private and Social Incentives for Fertility: Israeli Puzzles," Working Paper 8984, National Bureau of Economic Research.

Manski, Charles, F. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60, 531–542.

Meck, M. (1971): "Problems and Prospects of Social Services in Kenya," Weltforum Verlag, Munchen, IFO Institut fur Wirtschaftsforschung Munchen Afrikastudien 69.

Middleton, J., and G. Kershaw (1965): *The Kikuyu and Kamba of Kenya: The Central Tribes of the North-Esatern Bantu*. International African Institute, London.

Montgomery, M. R., G. Kiros, D. Agyeman, J. B. Casterline, P. Aglobitse, and P. C. Hewett (2001): "Social Networks and Contraceptive Dynamics in Southern Ghana," Population Council Research Division Working Paper No. 153, New York.

Morgan, W. T. W., and N. M. Shaffer (1966): *Population of Kenya, Density and Distribution*. Oxford University Press, Nairobi.

Moulton, B. R. (1986): "Random Group Effects and the Precision of Regression Estimates," *Journal of Econometrics*, 32, 385–397.

Mullahy, J. (1986): "Specification and Testing of Some Modified Count Data Models," *Journal of Econometrics*, 33, 341–365.

Mundlak, Y. (1978): "On the Pooling of Time Series and Cross Section Data," *Econometrica*, 46, 69–85.

Nauck, B. (1995): "Regionale Millieus Von Familien in Deutschland Nach der Politischen Vereinigung," in *Familie im Brennpunkt, Neuwid*, ed. by B. Nauck, and C. Onnen-Isemann, pp. 91–121.

Ominde, S. H. (1968): *Land and Population Movements in Kenya.* Heinemann, London.

Poston, D. L., and S. L. Mckibben (2003): "Using Zero-Inflated Count Regression Models to Estimate the Fertility of US Women," *Journal of Modern Applied Statistical Methods*, 2(2), 371–379.

Sander, W. (1995): *The Catholic Family.* Westview Press, Boulder.

Srinivas, M. N. (1994): *The Dominant Caste and Other Essays.* Oxford University Press, Oxford, Revised and enlarged edition.

Tanaka, K. (2000): "Medical Anthropological Study in Western Kenya and its Implications for Community Health Development," International Development Center of Japan Working Paper No. 55.

Throup, D. W. (2001): "Kenya: Revolution, Relapse or Reform?," Africa Notes, Centre for Strategic and International Studies, November, pp. 1-9.

Vuong, Q. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57, 307–334.

Watkins, S. C., N. Rutenberg, and S. Green (1995): "Diffusion and Debate: Controversy and Reproductive Change in Nyanza Province, Kenya," Paper presented at the Annual Meetings of the Population Association in America, San Francisco, April 6-8.

Were, G. S. (1967): *A History of the Abaluyia of Western Kenya, 1500-1900.* East African Publishing House, Nairobi.

Willis, R. J. (1973): "A New Approach to the Economic Study of Fertility Behaviour," *Journal of Political Economy*, 81(2 pt 2), 14–64.

Table 1: Children Ever Born (CEB) by Region and Ethnicity, Kenya DHS Sample 1993

	Region								
	Nairobi	Central	Coast	Eastern	Nyanza	Rift Valley	Western	μ^E_{CEB}	σ^E_{CEB}
Ethnicity									
Kalenjin	1.750	8.000	9.000	0.333	4.600	3.303	2.933	3.296	3.126
Kamba	1.688	3.714	2.619	2.842	2.000	3.800	-	2.729	2.799
Kikuyu	1.682	2.498	1.735	2.091	2.000	2.983	5.333	2.491	2.438
Kisii	1.308	1.667	4.000	0.000	2.716	2.625	8.500	2.687	2.803
Luhya	1.967	2.100	2.135	3.000	3.538	3.352	3.272	3.176	3.078
Luo	1.862	3.100	1.958	2.500	3.745	3.091	3.114	3.469	3.333
Meru/Embu	0.438	2.667	1.600	2.880	1.333	0.800	-	2.744	2.705
Miji/Swa	-	0.500	3.078	1.500	-	-	1.000	3.062	2.967
μ^R_{CEB}	1.692	2.512	2.865	2.829	3.310	3.224	3.271	2.983	-
σ^R_{CEB}	1.833	2.405	2.840	2.816	3.232	3.048	3.168	-	-

μ^E_{CEB} (σ^E_{CEB}) denote, respectively, mean (variance) of CEB by ethnicity.

μ^R_{CEB} (σ^R_{CEB}) denote, respectively, mean (variance) of CEB by region.

$-$ denotes a zero cell count. See Table 5.

Table 2: CEB by Region and Ethnicity for Various Age Groups

CEB	Age Group (years)								
	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-59	ALL
Region									
Nairobi	0.11	0.93	1.81	2.38	3.51	3.95	5.12	4.49	1.69
Central	0.17	1.03	1.93	3.28	4.38	5.63	6.33	5.92	2.51
Coast	0.31	1.28	2.57	3.87	5.17	6.34	6.78	6.53	2.86
Eastern	0.15	1.20	2.61	4.02	5.00	6.31	6.75	6.51	2.83
Nyanza	0.24	1.46	3.19	4.99	5.97	7.32	7.59	7.46	3.31‘
Rift Valley	0.27	1.38	3.12	4.54	5.90	6.95	7.71	7.25	3.22
Western	0.20	1.44	3.00	4.82	6.09	7.13	7.35	7.23	3.27
Ethnicity									
Kalenjin	0.27	1.39	3.36	4.90	6.12	7.27	7.80	7.48	3.30
Kamba	0.17	1.21	2.62	3.95	4.88	6.14	6.60	6.34	2.73
Kikuyu	0.18	1.00	1.93	3.27	4.38	5.60	6.34	5.91	2.49
Kisii	0.09	1.15	2.89	4.20	5.49	6.67	7.67	6.99	2.69
Luhya	0.21	1.40	2.93	4.58	5.70	6.91	7.39	7.10	3.18
Luo	0.36	1.67	3.11	4.85	6.25	7.51	7.60	7.56	3.47
Meru/Embu	0.10	1.05	2.41	3.64	4.90	6.25	6.84	6.52	2.74
Miji/Swa	0.30	1.35	2.80	4.24	5.38	6.60	6.77	6.68	3.06
ALL Regions/Ethnicities	0.22	1.29	2.75	4.22	5.44	6.61	7.11	6.83	2.98

Table 3: Mean Education (Years) by Region and Ethnicity.

	Region							
	Nairobi	Central	Coast	Eastern	Nyanza	Rift Valley	Western	μ^E_{ED}
Ethnicity								
Kalenjin	13.5	0.0	7.0	7.3	7.0	6.5	7.0	6.5
Kamba	9.2	4.9	6.9	6.7	8.5	6.6	-	6.9
Kikuyu	9.5	7.9	9.6	8.7	7.7	7.4	6.7	8.1
Kisii	10.5	6.0	5.5	12.0	6.8	6.7	6.0	6.9
Luhya	8.7	6.8	8.4	9.6	6.0	6.0	7.1	7.0
Luo	8.7	6.7	8.2	8.8	5.9	6.9	8.1	6.3
Meru/Embu	8.9	8.2	7.2	6.5	10.0	8.6	-	6.6
Miji/Swa	-	11.0	5.1	4.0	-	-	6.0	4.2
μ^R_{ED}	9.2	7.8	5.1	6.7	6.3	6.6	7.1	6.7

Note: μ^R_{ED} (μ^E_{ED}) denotes mean years of education by region (ethnicity).

$-$ denotes zero cell count.

Table 4: Distribution of the Sample by Region and Ethnicity

	Region							
	Nairobi	Central	Coast	Eastern	Nyanza	Rift Valley	Western	ALL
Ethnicity								
Kalenjin	4	1	1	3	5	1272	30	1316
	(1.0)	(0.1)	(0.1)	(0.3)	(0.4)	(67.6)	(3.6)	(18.1)
Kamba	80	7	84	672	2	10	0	855
	(20.7)	(0.9)	(10.0)	(57.3)	(0.1)	(0.5)	(0.0)	(11.7)
Kikuyu	154	741	34	22	3	298	3	1255
	(39.9)	(94.6)	(4.0)	(1.9)	(0.2)	(15.8)	(0.4)	(17.2)
Kisii	13	3	2	2	559	64	2	645
	(3.4)	(0.4)	(0.2)	(0.2)	(40.7)	(3.4)	(0.2)	(8.9)
Luhya	61	10	37	7	52	179	771	1117
	(15.8)	(1.3)	(4.4)	(0.6)	(3.8)	(9.5)	(91.6)	(15.3)
Luo	58	10	48	4	749	55	35	959
	(15.0)	(1.3)	(5.7)	(0.3)	(54.6)	(2.9)	(4.2)	(13.2)
Meru/Embu	16	9	10	460	3	5	0	503
	(4.1)	(1.1)	(1.2)	(39.2)	(0.2)	(0.3)	(0.0)	(6.9)
Miji/Swa	0	2	628	2	0	0	1	633
	(0.0)	(0.3)	(74.4)	(0.2)	(0.0)	(0.0)	(0.1)	(8.7)
ALL	386	783	844	1172	1373	1883	842	7283
	(100.0)	(100.0)	(100.0)	(100.0)	(100.0)	(100.0)	(100.0)	(100.0)

Table 5: Ideal Number of Children. Unwanted Fertility

	Age Group (years)								
	15-19	20-24	25-29	30-34	35-39	40-44	45-49	All Ages	UWF All Ages
Region									
Nairobi	2.85	2.65	2.94	3.00	3.47	3.21	3.69	2.95	-1.26
Central	2.65	2.60	2.84	3.28	3.67	3.60	3.98	3.06	-0.56
Coast	3.68	3.92	4.02	4.35	5.24	5.59	6.38	4.36	-1.69
Eastern	3.39	3.24	3.30	3.22	3.73	4.13	4.16	3.49	-0.66
Nyanza	3.54	3.67	4.06	4.53	4.71	5.47	5.52	4.17	-1.00
Rift Valley	3.89	3.65	3.76	4.18	4.76	4.81	5.22	4.1	-0.93
Western	3.81	3.59	3.85	4.36	4.65	5.63	4.80	4.15	-1.05
Ethnicity									
Kalenjin	4.04	3.75	3.96	4.40	4.81	4.99	5.36	4.24	-0.98
Kamba	3.34	3.31	3.50	3.35	4.09	4.08	4.48	3.6	-0.86
Kikuyu	2.77	2.65	2.89	3.26	3.92	4.06	4.18	3.16	-0.73
Kisii	3.29	3.39	3.92	4.29	4.38	5.28	4.71	3.96	-1.17
Luhya	3.73	3.65	3.79	4.23	4.41	5.19	4.83	4.05	-1.01
Luo	3.67	3.68	3.90	4.31	4.91	5.22	5.66	4.18	-0.89
Meru/Embu	3.29	3.02	2.86	3.03	3.41	4.02	3.92	3.27	-0.51
Miji/Swa	3.96	4.24	4.32	5.10	5.45	6.03	6.49	4.72	-1.90
All Regions/Ethnicities	3.54	3.44	3.62	3.93	4.43	4.78	4.94	3.87	-0.97

UWF - unwanted fertility.

Table 6: Children Ever Born.

Total Children ever Born	Frequency	Percentage
0	1568	26.16
1	777	12.96
2	693	11.56
3	563	9.39
4	540	9.01
5	436	7.27
6	440	7.34
7	309	5.16
8	246	4.10
9	196	3.27
10	128	2.14
11	49	0.82
12	34	0.57
13	10	0.17
14	3	0.05
15	2	0.03

Table 7: Parameter Estimates: Baseline Poisson Model

	$\widehat{\beta}$	$e^{\widehat{\beta}}$	p-value
Central	-0.174	0.840	0.001
Coast	-0.025	0.975	0.743
Eastern	-0.057	0.945	0.429
Rift Valley	-0.009	0.991	0.791
Nyanza	0.003	1.003	0.952
Cluster Level Controls			
Media access	-0.008	0.925	0.180
Water (nat. source)	-0.114	0.892	0.423
Water (piped in HH)	-0.235	0.791	0.094
Water (public source)	-0.177	0.837	0.245
Electricity (younger)	0.001	1.002	0.970
Electricity (older)	-0.022	0.978	0.643
Household Level			
Iron roof	-0.049	0.953	0.004
Tiled roof	-0.266	0.767	0.023
Other roof	0.011	1.011	0.932
Wood floor	-0.069	0.933	0.362
Tiled floor	-0.073	0.930	0.632
Cement floor	-0.049	0.952	0.020
Has radio	0.015	1.015	0.348
Has TV	-0.084	0.919	0.003
Has telephone	-0.044	0.957	0.378
Has bicycle	-0.016	0.984	0.322

Table 7 continued

	$\widehat{\beta}$	$e^{\widehat{\beta}}$	p-value
Individual Level			
Age	0.389	1.476	0.000
Age^2	-0.005	0.995	0.000
Primary Education (younger)	0.026	1.026	0.307
Primary Education (older)	0.042	1.043	0.137
Secondary Education (younger)	-0.245	0.783	0.000
Secondary Education (older)	-0.040	0.961	0.303
Higher Education (younger)	-0.456	0.634	0.000
Higher Education (older)	-0.224	0.799	0.002
Listens to radio (younger)	-0.003	0.996	0.680
Listens to radio (older)	-0.010	0.990	0.646
Resident since 1993	-0.038	0.963	0.094
Ethnic Effects			
Kalenjin	0.097	1.102	0.024
Kamba	0.006	1.006	0.927
Kisii	-0.021	0.979	0.731
Luhya	0.095	1.100	0.054
Meru Embu	-0.005	0.994	0.939
Miji_swah	0.030	1.030	0.730
Luo	0.139	1.150	0.020
-Ln L = 9607.8			

Table 8: Parameter Estimates: Zero Inflated Poisson Model with Social Interactions

	$\widehat{\beta}$	$e^{\widehat{\beta}}$	p-value
Central	0.279	0.757	0.000
Coast	-0.175	0.840	0.001
Eastern	-0.207	0.813	0.000
Rift Valley	-0.048	0.953	0.077
Nyanza	-0.013	0.987	0.670
Endogenous Effects: Cluster-level $\bar{C}_{e,cl} - C_i$			
Kalenjin (younger)	0.065	1.067	0.000
Kalenjin (older)	0.012	1.012	0.248
Kamba (younger)	0.068	1.071	0.000
Kamba (older)	0.001	1.001	0.935
Kikuyu (younger)	0.100	1.105	0.000
Kikuyu (older)	0.010	1.010	0.435
Kisii (younger)	0.042	1.043	0.012
Kisii (older)	0.034	1.034	0.005
Meru-embu (younger)	0.044	1.045	0.047
Meru-embu (older)	0.002	1.002	0.874
Mijikanda-Swahili (younger)	0.050	1.051	0.064
Mijikanda-Swahili (older)	0.051	1.052	0.014
Luhya (younger)	0.060	1.062	0.000
Luhya (older)	0.014	1.014	0.015
Luo (younger)	0.080	1.084	0.000
Luo (older)	0.001	1.001	0.905
Exogenous Effects: Cluster-level $\bar{E}_{e,cl}$			
Kalenjin (younger)	0.657	1.928	0.189
Kalenjin (older)	0.470	1.600	0.517
Kamba (younger)	0.187	1.205	0.148
Kamba (older)	0.017	1.017	0.928
Kikuyu (younger)	-2.250	0.105	0.001
Kikuyu (older)	-0.980	0.375	0.212
Kisii (younger)	-1.670	0.189	0.289
Kisii (older)	-0.158	0.853	0.800
Meru-embu (younger)	-2.083	0.124	0.100
Meru-embu (older)	0.529	1.697	0.385
Mijikanda-Swahili (younger)	0.251	1.285	0.571
Mijikanda-Swahili (older)	0.569	1.767	0.109
Luhya (younger)	-0.465	0.628	0.297
Luhya (older)	-0.605	0.546	0.339
Luo (younger)	0.777	2.174	0.229
Luo (older)	-2.009	0.134	0.132

Table 8: (continued)

	$\widehat{\beta}$	$e^{\widehat{\beta}}$	p-value
Cluster-level Controls			
Media accesss (radio)	-0.114	0.892	0.039
Electricity (younger)	0.051	1.052	0.277
Electricity (older)	-0.017	0.983	0.641
Water (nat. source)	-0.111	0.895	0.422
Water (piped into HH)	-0.172	0.842	0.213
Water (public source)	-0.100	0.905	0.493
Household Level			
Polygyny (younger)	-0.024	0.976	0.481
Polygyny (older)	-0.031	0.970	0.344
Usual Resident (younger)	0.049	1.051	0.297
Usual Resident (older)	-0.113	0.893	0.014
Iron roof	-0.058	0.943	0.000
Tiled roof	-0.330	0.719	0.001
Other roof	0.016	1.016	0.913
Wood floor	-0.090	0.914	0.206
Tiled floor	-0.016	0.984	0.891
Cement floor	-0.050	0.951	0.011
Has radio	0.012	1.012	0.425
Has TV	-0.091	0.914	0.001
Has telephone	-0.047	0.954	0.283
Has bicycle	-0.018	0.983	0.261
Individual Level			
Age	0.448	1.567	0.000
Age^2	-0.005	0.994	0.000
Primary Education (younger)	-0.058	0.943	0.096
Primary Education (older)	0.014	1.014	0.545
Secondary Education (younger)	-0.477	0.620	0.000
Secondary Education (older)	-0.095	0.909	0.001
Higher Education (younger)	-0.674	0.509	0.000
Higher Education (older)	-0.327	0.722	0.001
Listens to radio (younger)	0.017	1.012	0.169
Listens to radio (older)	-0.014	0.986	0.158
Resident since 1993 (younger)	-0.012	0.988	0.626
Resident since 1993 (older)	-0.108	0.898	0.020

Table 8: (continued)

	$\widehat{\beta}$	$e^{\widehat{\beta}}$	p-value
Inflation Sub-Model			
Kalenjin	0.932	2.541	0.434
Kamba	0.715	2.045	0.580
Kisii	-10.130	0.000	0.000
Luhya	1.286	3.618	0.272
Meru-embu	-10.416	0.000	0.000
Miji-Swah	0.665	1.944	0.653
Luo	2.084	8.040	0.068
-Ln L = 9453.4			
Vuong Test Statistic 3.60	(p = 0.000)		