

Multiple Social Interactions and Reproductive
Externalities: an Investigation of Fertility
Behaviour in Kenya

Sriya Iyer and Melvyn Weeks

November 2004

CWPE 0461

Not to be quoted without permission

Multiple social interactions and reproductive externalities: An investigation of fertility behaviour in Kenya

Sriya Iyer

Faculty of Economics & Politics
University of Cambridge

Melvyn Weeks*

Faculty of Economics & Politics
University of Cambridge

August 18, 2004

Abstract

This paper examines, theoretically and empirically, the impact of reproductive externalities on fertility behaviour in Kenya. We examine this issue by quantifying the effects of group membership on the number of children ever born. The focus of this study is the identification of structural forms of social interaction operating across individuals in the context of fertility behaviour. While structural forms of dependence may be separated from residual dependence, we also highlight the importance of different expressions of structural dependence which includes multiple expressions of social interaction. Using this idea of multiple social interactions, we use the 1998 Demographic and Health Survey data on 5994 women from Kenya to examine whether the 'local' effect of household-level influences and cluster-level residential settlement is important relative to the more 'global' effect of ethnicity on fertility behaviour. In so doing, we conclude that the importance of multiple social interactions is that if the assumption of a single model of interaction is made, erroneously, then it is possible to arrive at incorrect inferences.

Acknowledgements. We are grateful to Chris Brown for excellent research assistance in the early stages of this research and for making the DHS data available to us. For other data and information, we acknowledge the helpful assistance of David Nalo, Director of Statistics at the Central Bureau of Statistics in the Ministry of Finance and Planning, Kenya. For helpful comments and discussions, we thank Partha Dasgupta, Timothy Guinnane, Andrew Harvey, Hashem Pesaran, and seminar participants at the Brookings Institution, Max Planck Institute in Jena, St Andrews, and the University of Cambridge.

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

The shift in society from a high to a low-fertility regime, as witnessed in historical European populations and in some East Asian developing economies over relatively short periods of time, has shown that this has significant repercussions for economic development and the quality of life. Both economists and demographers have examined the balance between economic and non-economic factors in orchestrating a fertility transition. Classic studies of fertility such as those reported by the Princeton Fertility Project highlighted that the transition to low fertility in historical European populations occurred in a variety of different 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 in developed and developing countries alike have emphasised the role of the social climate and the influence of social interactions on demographic behaviour (Egero and Hammraskjold (1994), Kohler (2001), Hank (2001)). These empirical findings have led many economists naturally to focus on modelling the influence of social interactions on contemporary fertility transitions (Manski (1993), Manski and Mayshar (2002)) and on the reproductive externalities and coordination failures associated with fertility behaviour *per se* ((Dasgupta (2000); Kohler (2001))

In keeping with this literature, this study examines social interactions in the context of fertility behaviour in Kenya. 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. 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 multiple 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, be it local or global.

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. However, we are also interested in the influence of reproductive externalities - the social environment and context within which the woman makes her fertility choices. 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 this context we investigate the impact of both fertility behaviour and characteristics of appropriately defined groups of other individuals, including household members, members of a village cluster and ethnic group. 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)). For ex-

ample, a number of studies¹ have made the distinction between *social learning* and *social influence*. In a developing society like Kenya, the latter may occur through the effect of ethnic allegiance; whereas learning may be mediated by the experience of living within smaller groups, such as extended households, or clusters of households in close proximity. Fertility behaviour may then be mediated by social interactions both *within* one or more of these groups, or by interactions across different groups. Decomposing the influence of reproductive externalities into household-level, cluster-level and ethnic group interactions, has implications for the manner in which we conceive of social interactions influencing fertility behaviour.

An important conclusion of the present study is that the existence of multiple channels of social interactions implies that the analyst needs to exercise caution in making inferences. Specifically, we argue that both local and global interactions will be significant for fertility behaviour and that ignoring the existence of one or the other may lead to spurious inferences. Other studies of fertility behaviour do consider the question of spurious inference, but they do so more in the context of the omitted variables that account for unobserved variation between groups.

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 for both spatial proximity and ethnicity. In sections 5 we consider the distinction between multiple levels of social interaction, and in section 6 we examine the question of identification. In section 7 we consider a number of non-linear models of fertility behaviour, and in section 8 we introduce the data. We discuss our results in section 9, and section 10 concludes.

2 Social interactions and fertility

The theoretical literature on economic models of social interactions investigates the influence of non-market institutions on market institutions (see for example, Glaeser and Scheinkman (1999)). Brock and Durlauf (2003) argue that there have been two approaches adopted to investigate social interactions. The first is the impact of social interactions within predetermined groups (Akerlof (1997); Brock and Durlauf (2000), Brock and Durlauf (2001)). This literature emphasises how interactions affect individual and group-level outcomes in a cross-section. A second strand of this literature focuses on how social interactions lead to group formation with particular emphasis on the growth of residential neighbourhoods. This literature has focused on questions of geographic proximity, ethnicity, and the kinds of groups that lead to social interactions (Borjas (1995); Ginther, Haverman, and Wolfe (2000); Conley and Topa (2002)). Our study is located within the first strand of the literature. We regard ethnicity and cluster residence as fixed; and we do not consider the question of endogenous group formation.

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

¹See, for example, Montgomery, Kiros, Agyeman, Casterline, Aglobitse, and Hewett (2001).

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 fundamentals², 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)).

More recent 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 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. These hypotheses, as outlined by Manski (1993), are *endogenous effects*: the propensity of an individual to act varies, *ceteris paribus*, with the actions of their reference group. There are *exogenous effects*: the propensity of an individual to behave in some way varies, *ceteris paribus*, with the exogenous, observed mean characteristics of the reference group. And there are *correlated effects*: individuals in the same group tend to behave similarly because they have similar individual characteristics or face similar institutional environments.

In a fertility context, there is an endogenous effect if, *ceteris paribus*, a woman's children ever born (CEB) tends to vary with 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 three hypotheses would have different 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

²Horst and Scheinkman (2003) 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.

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)). The latter is responsible for prompting the less educated women, for example, also to adopt low fertility norms, which in turn causes eventually the society as a whole to move to a low-fertility equilibrium.

3 Overview of ethnic groups in Kenya

Multiple 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 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

³ According to the most recent Census.

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

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 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 division⁷ (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

⁵The 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.

⁶A 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)

⁷This was altered with the land reforms of 1950 when ownership of land was transferred to individuals.

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-grading⁸ (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). The Luhya do not have as powerful a clan organisation as some of the other groups (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

⁸The 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).

(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 heterogeneous 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.

⁹For 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)).

4 A Model of Reproductive Externalities with Geography and Ethnicity

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)) for household h

$$U_i = U(n_i, \mathbf{z}_i, \mathbf{x}_i, \varepsilon_i), \quad (1)$$

where U_i denotes the utility function of household h , 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.

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

$$V_i = (U_i(n_i, \mathbf{z}_i, \mathbf{x}_i); S^j(n_i, \mathbf{n}_{-i}^j), S^s(n_i, \mathbf{n}_{-i}^s), \varepsilon_i), \quad (2)$$

where V_i , total individual utility, is comprised of a private utility term $U_i()$, 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 members of j ; $\mathbf{n}_{-i}^j = (n_1^j, n_2^j, \dots, n_{i-1}^j, n_{i+1}^j, \dots, n_{M_j}^j)$, 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 $\mathbf{n}_{-i}^s = (n_1^s, n_2^s, \dots, n_{i-1}^s, n_{i+1}^s, \dots, n_{M_s}^s)$. 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. Rewriting (2) as additive in individual and social utility terms, we have

$$V_i = U_i(n_i, \mathbf{z}_i, \mathbf{x}_i) + S^j(n_i, \mathbf{n}_{-i}^j) + S^s(n_i, \mathbf{n}_{-i}^s) + \varepsilon_i. \quad (3)$$

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* interactions, whereby individual behaviour is influenced by the decisions of a relatively small number of well defined neighbours within location j ; 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

¹⁰Note that here within the household, we consider husband's and wife's preferences to be coincident and that we are not looking at 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.

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)). 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. We note that in (2) we have written the social utility terms as solely dependent on individual and group choices. We can also extend this function to include group level attributes in order to mitigate against spurious inference on the social interaction effects, as discussed further in section 7.2.

In focussing on that part of total utility attributable to social utility functions $S^j(\cdot)$ and $S^s(\cdot)$, it is obvious from (3) that cross-partial effects will be key objects. Differentiating the Beckerian first-order conditions with respect to social interactions yields the second-order conditions

$$\kappa_{il}^j = \partial^2 S^j(n_i, \mathbf{n}_{-i}^j) / \partial n_i \partial n_l^j > 0, \quad (4)$$

$$\kappa_{il}^s = \partial^2 S^s(n_i, \mathbf{n}_{-i}^s) / \partial n_i \partial n_l^s > 0, \quad (5)$$

where κ_{il}^j represents a measure of the disutility accruing to i from deviating from the behaviour of l . In examining (4) and (5) 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

$$S_1^j(n_i, \bar{n}_{-i}^j) = \theta_1 n_i \bar{n}_{-i}^j \quad (6)$$

$$S_2^j(n_i, \bar{n}_{-i}^j) = \frac{-\theta_2}{2} (n_i - \bar{n}_{-i}^j)^2. \quad (7)$$

(6) represents a proportional spillover form of dependence in the sense that for $\partial S_2^j(\cdot)/\partial n_i = \theta_1 \bar{n}_{-i}^j$, strategic complementarities are solely dependent upon the mean level of children ever born within location j . (7) represents social utility as a measure of conformism with $\partial S_2(\cdot)/\partial n_i = \theta_2(\bar{n}_{-i}^j - n_i)$. Under (7) 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

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

such that strategic complementarities are captured by a single parameter in both cases, based upon the discussion in section 3 on the varied characteristics of the ethnic groups, we can allow for this measure to be different across the groups suggesting different magnitudes of strategic complementarities across them (Bernheim (1994)).

5 Considering Multiple Levels of Social Interaction

Commentators such as Manski (2000) and Brock and Durlauf (2000) have been careful to point out the conditions under which a single form of structural dependence may be separated from residual (or correlated) dependence. In this paper we highlight the importance of different expressions of structural dependence. Below we examine what we believe to be three fundamental issues that need to be addressed in considering multiple levels of social interaction. These are first, understanding why social interactions need to be viewed as distinct processes; second, considering the extent or reach of social interactions; and third, elaborating upon the appropriate reference groups to measure the importance of social interactions empirically.

5.1 Distinct Processes: Social learning and social influence

In seeking to disentangle individual effects from the various manifestations of social interaction effects, it is necessary to go beyond the rather uninformative nomenclature of 'social interaction', and understand the precise mechanisms which translate aggregate fertility behaviour into an effect operating at the level of individual decision making. We consider two distinct processes by which social interactions influence fertility. First, through what we denote as a process of *social learning*,¹² which we use in the same sense as Ellison and Fudenberg (1993), namely where fertility decisions are determined, in part, by the accumulated experiences of others living locally. Second, *social influence* relates to the constraints on decision-making imparted by institutions, ethnic and/or political groups. In the context of Kenya,

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.

¹² Following Glaeser and Scheinkman (1999), we consider a typology of social interaction processes which includes multiple expressions of social interaction.

as also discussed in section 3, factors such as ethnicity play an important role in politics even today.

As an example, consider the case of social interaction and fertility. A process of interaction through social learning emphasises that women share experiences on the effectiveness of a new contraception technology or maternal and child health services, with greater use resulting in better information and reduced cost of use. In contrast an interaction through *social influence* highlights the effects of use by other community or ethnic group members in signalling the acceptance and lack of sanctions attached to the use of contraception or other health services.

5.2 The Extent of Social Interactions: Local and Global Effects

We believe the distinction between local and global interaction is useful in understanding the processes of social learning and influence. A local interaction occurs in relatively small groups consisting of individuals that may know each other and are able to observe individual fertility outcomes. 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. In contrast, global interactions, occurring amongst geographically dispersed ethnic groups, take the form of allegiances to group norms with the existence of sanctions for certain types of behaviour.

5.3 Measuring Social Interactions: The Choice of Reference Groups

At the outset we admit uncertainty as to what constitutes an appropriate reference group with respect to fertility decisions, in that our data contain no information which points to the existence of specific groups of individuals which, for example, represent a network within which social interaction takes place (see Montgomery, Kiros, Agyeman, Casterline, Aglobitse, and Hewett (2001)). In section 3 using a number of anthropological sources we identified that within rural populations in Kenya, families are grouped mainly in homesteads that are located in relatively small clusters, and this group forms the basis of community interaction. 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.¹³

In the context of fertility 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. Given that the Kikuyu exhibit the greatest sense of ethnic identity, this suggests that individuals departing from a set

¹³In 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)).

of group norms may feel a higher level of sanction, relative to members of another group. For these reasons, ethnic groups are especially likely to adopt and maintain different practices to those of other groups.¹⁴ 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. 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. In contrast, the household in Kenya constitutes a well-defined and fixed unit of analysis. It is important to understand here that individuals do not self-select into these groupings. They do not self-select at the level of the household. At the level of the cluster there is self-selection insofar as individuals locate in a cluster due to their ethnicity, although we control for this effect by conditioning our analysis on ethnicity.

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 $C_i \in \mathfrak{c} \subseteq \{0, 1, 2, \dots\}$ 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. This follows given that for any process bounded at zero, the variance will be an increasing function of the mean. In section 7.1 we examine a baseline Poisson model for the counts of children ever born, and determine the appropriateness of the attendant distributional assumptions. In addition although the failure to account for the group structure can result in inefficient estimation,¹⁵ the principal rationale for examining group effects is that the relationship between observations in the same groups is of interest. 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. In section 7.2 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.

¹⁴There has been a great deal of research on the effects on fertility of religion and ethnic group membership (Gellner (1981); Sander (1995); Iyer (2002)).

¹⁵See Moulton (1986), Chamberlain (1982) and Mundlak (1982).

6.1 A Baseline Poisson Model

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

$$f(C_i|\mathbf{x}_i) = \frac{e^{-\mu_i} \mu_i^{C_i}}{C_i!}, \quad C_i = 0, 1, 2, \dots \quad (8)$$

Given the nonnegativity constraint, the Poisson density in (8) 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. In particular, consider the defining property of the baseline Poisson model, namely the equidispersion property whereby $E(C_i|\mathbf{x}_i) = Var(C_i|\mathbf{x}_i) = \mu_i$. 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. In seeking to address this particular violation, we consider a specification which accommodates a general form of unobserved heterogeneity, and the preponderance of excess zeros. We also consider a specification which allows us to accommodate the fact that for younger women in the sample, we are observing truncated fertility lifetimes in that these women would not have completed their total childbearing at the time the survey was conducted.

6.1.1 Conditional Independence

A violation of the assumption of conditional independence in the Poisson model may occur as a result of a number of factors. Obviously if a count process is indexed by time, then the attendant immutable ordering suggests a number of ways to account for the dependence across observations. Within a cross-section departures from the standard exchangeability assumption may originate as a result of various contagion processes, where in general, there is greater uncertainty as to the appropriate ordering. The key issue here is whether contagion is viewed as a nuisance process, or is integral to the substantive focus of the research.¹⁶ In this study our objective is to construct a structural model of the contagion process 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 clusters of households. 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 . We note that after extending the Poisson conditional mean to reflect the violation of independence which we attribute to different expression of interactions across women in our sample, and the excess zeros problem (discussed further in section 7.1.3), it may still be necessary to control for residual dependence.

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

6.1.2 Neglected Heterogeneity

The equidispersion property restricts the variance of the observed CEB to be a known deterministic function of the vector of explanatory variables. If this restriction is not supported by the data invalid inference can result. As noted above, the variance of CEB exceeds the mean such that $Var(C_i|\mathbf{x}_i) = \phi\mu_i$, for $\phi > 1$. Given the form of the variance matrix of the Poisson Maximum Likelihood (ML) estimator, the estimated t-statistics will be inflated by a factor of $\sqrt{\phi}$. Dependent upon the value of ϕ , in cases of significant over dispersion, the ML Poisson estimates will be inefficient resulting in an increased tendency to make type two errors.

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)). As a consequence we utilise the PRM as a convenient baseline model. We consider two possible departures from this baseline. First, we extend the model in order to accommodate what we refer to as *general* unobserved heterogeneity. In this model the conditional mean is allowed 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

$$\begin{aligned}\mu_i(\omega_i) &= \exp(\gamma + \mathbf{x}_i'\boldsymbol{\beta} + \omega_i) \\ &= e^{\mathbf{x}_i'\boldsymbol{\beta}} e^{(\gamma + \omega_i)},\end{aligned}\tag{9}$$

noting that since ω_i is unobserved, (9) suggests a random-effects interpretation of this particular extension. Conditional on ω_i , the CEB count C_i is distributed Poisson

$$f(C_i|\mathbf{x}_i, \omega_i) = \frac{e^{-\mu_i(\omega_i)} (\mu_i(\omega_i))^{C_i}}{C_i!}, \quad C_i = 0, 1, 2, \dots,\tag{10}$$

with the unconditional probability given by

$$f(C_i|\mathbf{x}_i) = \int_0^\infty [e^{-\mu_i(\omega_i)} (\mu_i(\omega_i))^{C_i} / C_i!] g(\varpi_i) d\varpi_i.\tag{11}$$

For convenience, namely to ensure a closed form expression for (11), a gamma distribution with mean 1 and variance α is chosen for ϖ_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 $Var(C_i|\mathbf{x}_i) = \mu_i(1 + \alpha\mu_i)$. A test for over-dispersion is a Wald test of whether α is different from zero.

6.1.3 Excess Zeros

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. Thus, 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 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.

There are a number of alternate approaches to deal with this particular mixture model. For example, 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.¹⁷ In contrast the mixture model, which we write generically as

$$C_i \sim \left\{ \begin{array}{ll} 0 & \text{with probability } p_i \\ \Delta(\mu_i) & \text{with probability } 1 - p_i \end{array} \right\},$$

where Δ denotes the kernel density,¹⁸ 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.

6.1.4 Truncated Fertility Lifetimes

In both (8) and (10) 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. However, given that our sample includes women between 15 and 49 years of age, it is necessary to account for the fact that for younger women we effectively observe truncated fertility lifetimes. Denoting exposure time by S_i , and focussing on the conditional mean for the PRM, then we may modify the conditional mean, writing $\mu_i = \exp(\mathbf{x}_i\boldsymbol{\beta} + \ln(S_i))$, which simply means adding $\ln(S_i)$ as a covariate and constraining the coefficient to be unity. An alternative method which is adopted in this study, is simply to add 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.

¹⁷See Mullahy (1986).

¹⁸The kernel density in this instance may be either Poisson or Negative Binomial.

6.2 Conditional Mean Specification

In section 6.1 we motivated a number of extensions of the baseline Poisson model designed to mitigate against the inconsistency of the standard error estimates in the face of two manifestations of overdispersion. We now 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

$$E(C|\mathbf{x}^I, \mathbf{h}, \mathbf{p}, \mathbf{e}) = \exp(\mathbf{e} + \beta^e E[f(C|i \in \mathbf{e})] + \beta^c E[f(C|i \in cl)] + \gamma \overline{ed}_{cl} + \mathbf{h}'\boldsymbol{\theta} + \mathbf{p}'\boldsymbol{\kappa} + \mathbf{x}^{II'}\boldsymbol{\eta}), \quad (12)$$

where \mathbf{e} denotes a categorical variable denoting ethnic affiliation, $E[f(C|i \in \mathbf{e})]$ represents some function $f()$ of expected number of children ever born at the level of the ethnic group, and $E[f(C|i \in cl)]$ represents some function $f()$ of the number of children ever born at the level of the cluster (cl). \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 γ are scalar parameters denoting, respectively, endogenous and exogenous effects.

6.2.1 Endogenous and Exogenous Interaction Effects

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

$$\beta^e E[f(C|\mathbf{e})] = \partial S_2^e(n_i, \bar{n}_{\mathbf{e}-i}) / \partial n_i = \beta^e (\bar{n}_{\mathbf{e}-i} - n_i), \quad (13)$$

where n_i denotes the number of children ever born to woman i , and $\bar{n}_{\mathbf{e}-i}$ is the mean CEB for ethnic group \mathbf{e} excluding i ; β^e represents the cross-partial effect¹⁹ $\frac{\partial^2 S_2^e(\cdot)}{\partial n_i \partial \bar{n}_{\mathbf{e}-i}}$. Although the conformist social utility function $S_2^e(\cdot)$ exhibits a strategic complementarity that is constant *within*²⁰ a given ethnic group, it is allowed to vary *across* groups given the recognition that ethnic groups may possess loss functions which differentially penalize behaviours which are different from the norm. In addition we also construct an endogenous interaction effect $E[f(C|i \in cl)]$ designed to capture conformists effects which are localised in the sense of being confined to women who reside in the same cluster of households. It is important to be clear as to the distinction between $E[f(C|i \in \mathbf{e})]$ and $E[f(C|i \in cl)]$. Whereas the former is constructed utilising the *global* ethnic population as the reference category, the latter utilises the *local* population resident within the village cluster. Note here that in many cases, the cluster populations are homogenous with respect to ethnic group.

Such a formulation allows us to examine, for example, whether after controlling for social interaction effects accruing through the desire to conform to a set of

¹⁹The notation $S_1()$ and $S_2()$ refers, respectively, to the proportionate and conformist social utility functions presented in Section 4.

²⁰We do not have information on the *distance* between individuals within the same ethnic group or cluster.

behaviours ascribed by ethnicity, there is a residual effect attributable to the proximity of neighbours. As a corollary of this, we are able to determine whether there is a local (geographic) expression of ethnicity in terms of interaction effects. By including both of these effects in the conditional mean we note that β^e (β^c) captures the endogenous effect after partialling out variance of all other factors which also includes the endogenous effect constructed at the level of the cluster (ethnic group).

We represent exogenous effects attributable to observing the education of other women in the same ethnic group and cluster using

$$\overline{ed}_{cl} = \frac{1}{n_{cl}} \sum_{i \in cl} ed_i \quad (14)$$

where ed_i is the education level of the i^{th} woman. 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 also examine whether such interactions may be manifest in the form of additional exogenous effects such as discussions about family planning. A mean level variable measured whether or not family planning had been discussed with friends and neighbours. This form of interaction is important if, for example, a woman was likely to discuss family planning with friends and neighbours, 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 friends and neighbours is important for fertility.

6.2.2 Cluster-level Controls

In section 4 we extended the atomistic utility model by including social utility functions with the objective of capturing that component of total utility due to CEB that we may attribute to the effects of the fertility choices of other women. 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. One key attribute that we need to control for 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 in turn reduces the demand for children, and hence the fertility rate (Iyer (2002); Aggarwala, Netanyahu, and Romano (2001)). Access to fuel at the cluster level 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.

6.2.3 Household-level Interactions

A number of mean-level effects were created to capture the effects of interactions within the household. One measure of interaction is a measure of polygyny: there were 30% of households who had more than one wife living in the household. Recent other studies have also examined the role of polygynous marriage norms in influencing demographic behaviour (Hogan and Biratu (2004)).²¹ The number of usual residents in the household, where usual residents excludes the number of children-ever-born to women, was also included. 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.

6.2.4 Household-level Controls

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. First, the quality of both the flooring of the home and the quality of the roof were used in the model. 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.

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. 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. Ownership of a bicycle is important for access to fuel and water.²²

6.2.5 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

²¹It is expected that households which had more than one wife would have lower fertility due to help provided with child-care. Alternatively, it may reduce fertility as information about contraception may be more discussed among co-residents in the household.

²²For example, other demographic studies have found that ownership of a bicycle considerably reduces the time spent collecting water for a households' daily requirements (Iyer 2002).

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 radio²³. A binary variable recorded how long the woman had lived in the community. 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. 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.

7 The Identification Problem

Below we consider the general problem of identification. The observed data consists of $(C, \epsilon, \mathbf{x})$ where C is scalar outcome measuring children ever born (CEB), ϵ is a categorical variable identifying ethnic group, \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. A linear model of CEB which considers different manifestations of social interaction is written as

$$C = \alpha + \beta E[f(C|i \in \epsilon)] + E(\mathbf{x}|\epsilon)' \boldsymbol{\gamma} + \mathbf{x}' \boldsymbol{\eta} + \mathbf{u}, \quad (15)$$

where α and β are unknown scalar parameters, and $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$ are, respectively, $k \times 1$ and $l \times 1$ parameter vectors, with $l \leq k$. $E(u|\epsilon, \mathbf{x}) = \epsilon' \boldsymbol{\delta}$.

The conditional mean of C given (ϵ, \mathbf{x}) may be written as

$$E(C|\epsilon, \mathbf{x}) = \alpha + \beta E[f(C|i \in \epsilon)] + E(\mathbf{x}|\epsilon)' \boldsymbol{\gamma} + \epsilon' \boldsymbol{\delta} + \mathbf{z}' \boldsymbol{\eta}. \quad (16)$$

Endogenous effects on individual fertility outcomes, in the sense that the mean children ever born (CEB) by ethnic group is a determinant of individual CEB are present for $\beta \neq 0$. For $\boldsymbol{\gamma} \neq 0$ we observe a *exogenous* effect whereby CEB varies with $E(\mathbf{x}|\epsilon)$, the mean characteristics by ethnicity. Focussing upon a single element of \mathbf{x} , say x_1 denoting education level, we are interested in capturing the effect of more highly-educated women in a population on fertility decisions *after* controlling for the level of education of the individual. We would expect that conditional upon ethnicity, CEB would be lower for more highly educated women than others, such that $\eta_1 < 0$. In addition, the characteristics of these more highly educated women may filter through to influence CEB such that $\gamma_1 < 0$ orchestrating a transition in fertility.

²³ Messages about this are routinely broadcast on the national radio channel, the 'Voice of Kenya' in English, Swahili and the vernaculars.

Correlated effects operate through δ and reflect the presence of common unobservable characteristics: these may range from factors such as the influence of the local religious clergy or women's perceptions about the costs and benefits of childbearing.

As it stands (16) is not estimable given the endogenous regressor $E(C|\epsilon)$. In order to obtain an expression for $E(C|\epsilon)$ we marginalise both sides of (16) with respect to \mathbf{x} , giving

$$E[f(C|i \in \epsilon)] = \alpha + \beta E[f(C|i \in \epsilon)] + E(\mathbf{x}|\epsilon)' \gamma + \epsilon' \delta + E(\mathbf{x}|\epsilon)' \eta. \quad (17)$$

For $\beta \neq 1$ we may rewrite (17) as

$$E[f(C|i \in \epsilon)] = \alpha/(1 - \beta) + E(\mathbf{z}|\epsilon)'(\gamma + \eta)/(1 - \beta) + \epsilon' \delta/(1 - \beta). \quad (18)$$

The genesis of the identification problem derives from the simple observation that in (15) we specify a model for children ever born which is linear in a mean endogenous effect, $E(C|\epsilon)$, and a mean exogenous effect, $E(\mathbf{x}|\epsilon)$. It therefore follows that the reduced form representation of the conditional mean $E(C|\epsilon)$ will be a function of $E(\mathbf{x}|\epsilon)$. 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 $E(C|\epsilon)$, $E(\mathbf{x}|\epsilon)$, and $E(u|\epsilon, \mathbf{x})$, and observe that, not surprisingly, the reduced form $E(C|\epsilon)$ is a function of $E(\mathbf{x}|\epsilon)$.

To determine exactly what is identified we insert (18) into (16) giving

$$E[f(C|i \in \epsilon), \mathbf{x}] = \alpha/(1 - \beta) + E(\mathbf{x}|\epsilon)'[(\gamma + \beta\eta)/(1 - \beta)] + \epsilon' \delta/(1 - \beta) + \mathbf{x}' \eta. \quad (19)$$

From (19) we see that the composite parameters $\alpha/(1 - \beta)$, $(\gamma + \beta\eta)/(1 - \beta)$, $\delta/(1 - \beta)$, together with η are identified conditional upon the regressors $[1, E(C|\epsilon), \epsilon, \mathbf{x}]$ being linearly independent in the population.

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 zero restrictions. For example, Brock and Durlauf (2000) achieve identification by locating an individual level variable whose group level average can be ruled out a priori as a contextual effect²⁴. A more general approach, and the one used in this study, is to apply a link function to the conditional expectation $E(C|x, \mathbf{x})$. Namely, instead of assuming that this expectation is linear we write

$$E(C|\epsilon, \mathbf{z}) = F(\alpha + \beta E(C|\epsilon) + E(\mathbf{x}|\epsilon)' \gamma + \epsilon' \delta + \mathbf{x}' \eta), \quad (20)$$

where $F(\cdot)$ is a nonlinear monotonic transformation which maps the conditional mean from \Re to the unit interval.

As a consequence of this mapping we note that although the conditional mean $E(C|\epsilon)$ will still be a function of $E(\mathbf{x}|\epsilon)$, the linear dependence between these two objects is automatically removed given that the support of $E(C|\epsilon)$ is bounded: between zero and one for a simple binary choice model of fertility, and between zero and ∞ ,

²⁴Note that the existence of such variables will increase the precision of the estimates in a nonlinear setting.

where \varkappa is the maximum number of children ever born in either a multinomial²⁵ or count model representation. Subsequently, as long as the support of $E(\mathbf{z}|x)$ is sufficient to reveal the nonlinear dependence between an endogenous and a contextual effect, Manski's reflection becomes blurred and identification is achieved. Further it is important to note that the same arguments for the identification of the cluster-level endogenous effects, $E[f(C|i \in \mathfrak{cl})]$ apply. Separate identification of these two effects follows from variation within the ethnic group and the cluster groups.

Although the observation that the introduction of a non-linear transformation facilitates the identification of an otherwise unidentified model is not new²⁶, it is important to note that in this instance identification is achieved as a consequence of a correctly specified functional form. Namely, in models of fertility choice the link function is required in order to correctly represent the nature of the observed data. Identification is therefore achieved as a result of the specified nonlinear model, rather than as a specifically targeted means to blur the linear dependence between endogenous and contextual effects.

8 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),²⁷ has data on the Kalenjin, the Kikuyu, Luhya, Luo, Kamba, Kisii, Mijikenda/Swahili and Meru/Embu.²⁸ 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.²⁹ 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

²⁵ See Brock and Durlauf (2003) for a discussion of multinomial choice with social interactions.

²⁶ For example, Heckman's selection model.

²⁷ 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.

²⁸ 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.

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

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 5 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 group³⁰.

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 6 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 6 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 com-

³⁰It is important to understand here that individuals do not self-select into these groupings. 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.

pleted 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 sterilisation 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.

9 Results

We present results from a baseline PRM model. We do this for two reasons. First, we note that although heterogeneity in the Poisson model over and above that which is consistent with the specification of the conditional mean implies overdispersion, heterogeneity is not implied by overdispersion. In this respect it is possible for the overdispersion inequality $Var(C_i|\mathbf{x}_i) > \mu_i$, for $\mu_i = E(C_i|\mathbf{x}_i)$, to hold, as it does in the case of count data CEB in Kenya, but where there is no unobserved heterogeneity. Taking the PRM as the null model, we first test for departures from equidispersion based upon a general form of unobserved heterogeneity as represented by the negative binomial model with variance $Var(C_i|\mathbf{x}_i) = \mu_i(1 + \alpha\mu_i)$. If we fail to reject the null, we then extend the PRM in the direction of the zero-inflated Poisson alternative; if we reject the null we take the Poisson-gamma mixture (Negative Binomial) as the new baseline model, and determine whether after for conditioning on the general form of heterogeneity, an excess zeros problem remains. Second, the PRM model provides a useful point of comparison with which to compare our parameter estimates obtained from a conditional mean specification which includes a number of interaction effects.

The results of the baseline PRM are shown in Table 7³¹. 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 number of children ever born due

³¹Standard errors are calculated using the Huber (1967) method to account for intra-cluster correlation.

to a unit change in the regressors.³² However, since the predicted count is given by $e^{\mathbf{x}'\hat{\beta}}$, we also report the factor change in the expected count for a unit increase in a given x_k , which we denote $e^{\hat{\beta}_{x_k}}$.

The results of the different model estimations are discussed in the following order: first, Table 7 presents the Baseline Poisson Model. Table 8 presents the Zero-Inflated Poisson model with cluster, household and individual-level effects included.

9.1 Baseline Poisson Model

The results from the estimation of the Baseline Poisson Model are shown in Table 7³³. In keeping with previous discussion about different levels of effects, this table distinguishes between, and depicts, the individual effects alongside the household-level and cluster-level controls. In this model we do not consider any interaction effects. To consider the effect of potential fertility determinants at different levels, we initially focus on a single covariate, access to the media. As a cluster-level control, access to the media as measured by the proportion of households in the cluster which had access to a radio, exerts a significant negative effect on children ever born; if the woman lived in a cluster which had a higher proportion of radios, then this reduced her expected CEB by approximately 10%. This variable is not significant at the household level. 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, is not significant.

At the household level the 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. For example, if a woman lived in a house which had an iron roof, the expected number of children ever born would be 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 20%. This is also reflected in the fertility of women who lived in homes which had cement floors (compared to sand), with on average 4% lower CEB. 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

³²This follows since

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

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

³³Note that in all models we included dummies to control for any residual regional effects (Western, Central, Coast, Eastern, Rift Valley, and Nyanza). Since all regional effects were insignificant, we choose to exclude these parameter estimates from our tabulated results.

to her younger counterpart. Education exerted a negative impact on fertility, but only at higher levels such as secondary and higher education: the completion of secondary education multiplies the expected number of children by a factor of 0.84; if she had completed higher education, then this factor falls to approximately 70%. In both cases the base category was women with no education. Primary education has the expected positive relationship with CEB. A woman's ideal number of children is also shown positively to increase the likelihood of greater CEB. Using the Kikuyu as the base group, the interpretation of the parameter estimates for the ethnic group effects, are as expected, and consistent with Table 1. Specifically, since the Kikuyu have, on average, the lowest CEB across all ethnic groups, we note that with the exception of the Mijikenda/Swahili group, all (relative) ethnic effects are positive. In addition the magnitude of these effects also reflect the relative CEB values, with, for example, the Luo having the highest (unconditional) CEB and the largest ethnic effect.

The Baseline Poisson model represents a valuable point of departure in demonstrating the importance of the individual-level factors. However, it does not take into account the effect of endogenous and exogenous interaction effects at the level of household, ethnicity and clusters. The model also maintains an equidispersion assumption and is not able to adequately model a preponderance of the zero counts. A likelihood ratio test ($H_0 : \alpha = 0$, $H_1 : \alpha > 0$) of whether the PRM model should be extended to account for a general form of unobserved heterogeneity was conducted and we were not able to reject the null at any reasonable significance level. We then extended the Poisson model to account for the specific form of unobserved heterogeneity - namely the fact that we do not observe an infertility indicator. Table 8 presents the results of a Zero-inflated Poisson regression estimated on the Kenyan data, with ethnic, cluster and household endogenous and exogenous effects, along with individual-level effects. Although the model contains a large number of variables based upon cluster aggregates, standard errors are again calculated using the Huber (1967) method to adjust for any residual intra-cluster correlation.

9.2 Zero-Inflated Poisson Model (ZIP)

The 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.

9.2.1 Endogenous effects

We differentiate between two types of endogenous effects: effects emanating from within the village cluster and effects which emanate within the broader ethnic group. The localised endogenous effects are captured by a parameter which weights the variable $(\bar{C}_d - C_i)$ - namely the difference between a mean-level and individual CEB for women who live in the same cluster. We observe that for all ethnic groups these parameters are insignificant, except for the Kalenjin for whom the impact is positive and for the Kikuyu for whom the impact is negative. For the Kikuyu, every additional unit increase in mean CEB decreases the number of children by 4%,

holding all other variables constant.

When endogenous effects are measured using the entire ethnic population as the reference group and measured using a conformist social utility function $(\bar{C}_e - C_i)$, all parameters were highly significant. For all ethnic groups with the exception of the Kikuyu, the estimated parameters were in the range 0.16-0.20, which implies a range of factor change effects of 1.18-1.22. For the Kikuyu, this effect was considerably greater ($\hat{\beta}^e = 0.26$, and $e^{\hat{\beta}^e} = 1.3$). In seeking to explain the Kikuyu outlier, we note that the Kikuyu have a highly evolved sense of ethnic identity, and that for this group, households are generally comprised of groups of individuals which are related by blood and marriage (the *mbari*). As a result one might expect a greater tendency for fertility decisions to reflect ethnic group norms.

The findings on the Kikuyu endogenous effects with respect to cluster and ethnicity need further reflection. While ethnicity seems to be important for fertility among the Kikuyu at a global level (in that ethnic norms might encourage childbearing among women of this community more generally), the effect at the level of the cluster works in the opposite direction. Although the cluster effect is relatively small and significant, it is dominated by the strong positive ethnicity endogenous effect. This could be because of the effect of another type of 'reproductive externality'. As discussed in section 3, lineage and clan considerations may be pronatalist: these work to increase fertility. On the other hand, in a more local setting, women observe the negative effects of population growth in terms of its impact on natural resources and access to other resources, and may then respond to any observed increase in mean CEB in their immediate vicinity. Studies of fertility in other developing societies have shown that rural women are acutely aware of - and indeed keenly express their views in household surveys on - the limitations of rapid population growth in their village or area (Iyer (2002): 213). Frequently these views are completely at odds with their *actual* fertility behaviour, creating an externality which makes attempts to formulate government policy to reduce population growth in these poor countries, very complex indeed. The fact that these seemingly opposing effects arise not only among rural South Indian populations, but are evident here among this Kenyan ethnic group as well, makes the findings on the Kikuyu endogenous effects all the more fascinating. Moreover, the Kikuyu women in the sample are, on average, the most highly educated group of women, so it may also be that they are more likely to be aware of the negative effects of high fertility, and are also more likely to show their responses to it compared to women of other ethnic groups. This argument would also be consistent with the positive endogenous effect found for the Kalenjin women, who are, on average, some of the least well-educated of the women in the sample.

In interpreting the overall significance of the global and local endogenous effects, there are some caveats that need to be acknowledged. First, that in models that are estimated empirically to capture these effects, the standard *ceteris paribus* condition with respect to all other controls holds. Second, that the size of the reference group differs substantially: at the level of the village cluster, the number of households are relatively small (on average 17 households). Therefore, if in addition the within cluster CEB exhibits low variance, then it will be difficult to pin down precisely the

magnitude of the endogenous interaction effects.

9.2.2 Exogenous effects

An exogenous education effect was significant for two ethnic groups - the Kamba (at the 7% level) and the Kisii (12%).³⁴ For a 1% increase in the proportion of Kamban women in a cluster who had completed secondary education, we observe a 37% decrease in Kamban fertility, which is a sizeable effect. For the Kisii ethnic group the effect was to decrease fertility by as much as 55%. We can explain such sizeable education effects: for Kenya as a whole, the Kamba have, on average similar rates of schooling attainment to other groups such as the Kalenjin. Yet these levels are still below the Kikuyu, for example. This might suggest that groups which have lower levels of education, on average, are also more likely to be influenced by social interactions derived from the better-educated members of their ethnic groups. This finding is also consistent with the observation that when the indirect effects are very significant, the multiplier is very large (Horst and Scheinkman (2003b)). What is perhaps the most interesting aspect of these findings on the exogenous education effects, is that they are strong even after controlling for the individual levels of educational attainment.

9.2.3 Cluster-level controls

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 mean-level cluster effect for access to electricity exerted a negative impact on CEB, reducing CEB by about 3%, but this was significant only at the 18% level. This finding is consistent with other economic analyses of poor countries which suggest that the provision of fuel infrastructure, such as electricity, is important for fertility decline as it reduces the reliance on children to collect firewood to meet the household's daily fuel consumption requirements (Dasgupta (1993)).

9.2.4 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 in order to capture characteristics at the household level so as to reduce the possibility that any inference on the significance of the household-level interaction effects is not confounded by household-level omitted variables. As far as the household effects are concerned, none of the household-level social effects were significant.

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

³⁴We note that the mean level variable here is the proportion of women in a cluster who have completed secondary education.

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 12%. This finding is unsurprising as roofing quality is a strong indicator of income in a rural household. This finding seems to be a robust result as it was also important in the baseline model. None of the floor quality variables were, however, significant. 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, as in the case of the PRM model, was insignificant. In contrast to the PRM model, television ownership is now insignificant. Another income indicator - rooms for sleeping - depicted negative effects on fertility at the 10% level.

9.2.5 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 positive effect on fertility consistent with other studies that a few years of education merely weakens existing taboos (such as not practising abstinence upon attaining grandmother status), and that this accounts for an initial increase in fertility. At the individual level, the education variables were also very significant and displayed the expected negative signs. If a woman was educated to the secondary level, she depicted on average 3% lower fertility; if she had had higher education, then this reduced her fertility by about 7%. If the 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 by approximately 2%. The finding that despite controlling for both ethnic and cluster level interactions, there are still significant ethnic effects is noteworthy.

9.2.6 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 with the following predictors. First, although we do not observe any reliable indicators of infertility, some women in the sample responded zero to the question 'What is your ideal family size?' Although, this may of course measure a true desire for not having children, in instances where the answer to this question was zero and the actual number was zero, we included a binary indicator which we believe represents a proxy for infertility. Current age was used to control for the fact that we are observing truncated fertility lifetimes for that portion of women who have not come to the end of their child bearing years. In both cases we observe the expected negative effect.

We also included ethnic dummies, again with Kikuyu as the base ethnic group. The rationale for inclusion of the ethnic dummies is that the existence of infertility may vary across the ethnic groups due to a number of unobserved factors that are

likely to be highly correlated with the incidence of infertility such as genetic composition, nutritional status, health status, education, and so forth. 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, with the exception of the Luo. This practice may lead to lower fertility if it endangers the health of women. The existence of this practice is consistent with the findings of the inflation sub-model in which the only group for which we observe a significant ethnic effect is the Luo which suggests that there are, on average, a smaller proportion of infertile women in the sample for this group. We also emphasise the strength of this effect, namely an estimated factor change effect of 0.076, which indicates that Luo women, (relative to the Kikuyu), have, on average, a 93% lower probability of never having children. As discussed in section 3, anthropological studies of the Luo have commented that this group especially exhibit a strong preference for large families (Watkins, Rutenberg, and Green (1995)). The combination of the findings from the inflation sub-model on the Luo, the anthropological evidence, and Table 1 which shows that the Luo depict the highest fertility of any ethnic group in Kenya, seems to suggest that their high fertility in this instance may be the outcome of a combination of factors including a predisposition towards relatively lower levels of infertility, compared to other 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 rejected³⁵.

10 Conclusion and policy relevance of research

In the debate about whether it is individual economic circumstances that influences fertility behaviour, or whether there are factors within the wider society that have implications for fertility, the role of social interactions has been observed to matter. This paper has also examined the importance of social interactions, in the context of fertility behaviour in Kenya and has stressed the importance of the multiplicity of social interactions.

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 strategic interactions. In this paper we argue that it is possible to identify the channels through which these interactions occur, and that they are plural rather than singular in nature. We argue that these channels operate at the household, cluster and ethnic level. This is an important finding of the research because it suggests that 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 conclusion of the paper is that in order to influence fertility in Kenya, we need to be aware of the importance of both local and global interactions in the

³⁵The value of the test statistic was 16.58, with a p-value of 0.000. See Vuong (1989) for details of the form of this test.

form of ethnicity and geography. The key finding is that once the influence of a 'global' form of interaction - ethnicity - is included, the influence of a more 'local' form of interaction - geographical clusters - becomes insignificant. This implies that we need to be cautious when specifying the channels of social interaction and in drawing inferences from them.

From a policy perspective, the analysis conducted here suggests that in Kenya, we will need to target the demand for children via influencing people's desires about fertility. It is important to recognise that economic characteristics interact with norms so that, for example, raising education levels for women of all ethnicities will be significant for fertility declines. For example, educating women is important as higher educated women depict lower fertility. 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.

Ethnicity is significant for fertility to the extent that norms of behaviour among members of the same ethnic group, may coincide. Relatedly, if ethnicity does matter, are there are other possible orderings which matter for fertility? For example, it is acknowledged that although the form of ordering in a time series is both singular and immutable, within a cross-section there may exist more than one ordering, and in addition, the relative importance of these orderings may change over time. In this study we have entertained three possible orderings: (i) that dependence exists within the household; (ii) that dependence exists across individuals that reside in a cluster of households; and (iii) that dependence exists across individuals within the same ethnic group. In this context a key issue is that 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 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 three orderings - the household, the cluster, and the ethnic group - is unlikely to change. This is because, as also argued in section 3, the ethnic norms of the different ethnic groups in Kenya have salient characteristics that are important for individuals across generations - these norms are, in all likelihood, immutable, and hence seem to 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. Geography in the form of localised cluster interactions, is an important kind of social interaction that influences fertility behaviour in Kenya, *but in most cases, when ethnicity is omitted from the model.*

The research presented in this paper has argued that incorporating social interactions allows us to extend the basic household demand model of fertility to capture that component of individual utility attributable to a fertility choice that is dependent upon the fertility choices of others. We modify the social utility function by

allowing the measure of dependence to vary across ethnic groups, and we add cluster level controls to mitigate against spurious inference. But it strikes us that further research in this area could motivate other extensions as well.

One line of inquiry may be to consider that endogenous and exogenous effects are not simply 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 externality, but as we see it this would be a potential avenue for future research in this area.

Lastly, the key message that emerges from this study is that since fertility is influenced both by 'local' interactions and by 'global' norms, from a policy perspective, a specifically targeted intervention in an area of high fertility, or towards specific groups in society, arguably will be more effective than a 'one size fits all' approach to population policy. Therefore, both theoretically and empirically, if we are interested in reducing fertility in poor societies such as Kenya, or indeed drawing lessons for other developing societies more widely, we think it apposite to sound a word of caution: *Social interactions matter, but multiple interactions matter even more*. Failure to recognise their importance and interdependence will lead only to misspecifications, and more worryingly for population policy in poor societies, the problem of spurious inference.

References

- (1970): *Population Census 1969* Republic of Kenya, Vol. 1, Nairobi.
- 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. KEKOVOLÉ (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.
- 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.
- BERNHEIM, B. D. (1994): "A Theory of Conformity," *Journal of Political Economy*, 102(4), 841–77.
- 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," University of Wisconsin mimeo.
- 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.
- 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.
- EGERO, B., AND M. HAMMRASKJOLD (eds.) (1994): *Understanding Reproductive Change: Kenya, Tamil Nadu, Punjab, Costa Rica* Lund. Lund University Press.
- ELLISON, G., AND D. FUNDENBERG (1993): "Rules of Thumb for Social Learning," *Journal of Political Economy*, 101(4).
- 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. SRUNGBONMEE (2003): "Culture and the Economy in Kenya," Paper submitted for the 1st Annual undergraduate research symposium, University of South Florida, Spring.
- 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 Contraceptive 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.
- 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. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60(3), 531–42.
- (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.
- MECK, M. (1971): "Problems and Prospects of Social Services in Kenya," Weltforum Verlag, Munchen, IFO Institut fur Wirtschaftsforschung Munchen Afrika-studien 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.
- 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

	Nairobi	Central	Coast	Eastern	Region Nyanza	Rift Valley	Western	μ_{CEB}^E	σ_{CEB}^E
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_{CEB}^R	1.692	2.512	2.865	2.829	3.310	3.224	3.271	2.983	-
σ_{CEB}^R	1.833	2.405	2.840	2.816	3.232	3.048	3.168	-	-

μ_{CEB}^E (σ_{CEB}^E) denote, respectively, mean (variance) of CEB by ethnicity.

μ_{CEB}^R (σ_{CEB}^R) 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)								ALL
	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-59	
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							μ_{ED}^E
	Nairobi	Central	Coast	Eastern	Nyanza	Rift Valley	Western	
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_{ED}^R	9.2	7.8	5.1	6.7	6.3	6.6	7.1	6.7

Note: μ_{ED}^R (μ_{ED}^E) denotes mean years of education by region (ethnicity).

— denotes zero cell count.

Table 4: 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 5: Distribution of the Sample by Region and Ethnicity

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

Table 6: Ideal Number of Children. Unwanted Fertility

	Age Group (years)							All Ages	UWF All Ages
	15-19	20-24	25-29	30-34	35-39	40-44	45-49		
<i>Region</i>									
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
<i>Ethnicity</i>									
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
<i>All Regions/Ethnicities</i>	3.54	3.44	3.62	3.93	4.43	4.78	4.94	3.87	-0.97

UWF - unwanted fertility.

Table 7: Children Ever Born: Poisson Model

	$\hat{\beta}$	$e^{\hat{\beta}}$	p-value
<i>Cluster Level Controls</i>			
Media access	-0.105	0.901	0.085
Electricity in HH	0.002	1.002	0.963
Water (nat. source)	0.001	1.000	0.996
Water (piped in HH)	-0.089	0.915	0.524
Water (public source)	-0.040	0.961	0.793
<i>Household Level</i>			
Iron roof	-0.053	0.948	0.002
Tiled roof	-0.234	0.792	0.049
Other roof	0.024	1.024	0.879
Wood floor	-0.062	0.940	0.406
Tiled floor	-0.030	0.970	0.777
Cement floor	-0.038	0.963	0.077
Toilet facility	-0.001	0.999	0.741
Has radio	0.012	1.012	0.474
Has TV	-0.087	0.916	0.003
Has telephone	-0.056	0.946	0.290
Has bicycle	-0.017	0.984	0.343
Rooms (sleep)	-0.004	0.996	0.526
<i>Individual Level</i>			
Age	0.383	1.467	0.000
Current Age	-0.005	0.995	0.000
Education (primary)	0.057	1.058	0.009
Education (secondary)	-0.174	0.840	0.000
Education (higher)	-0.359	0.699	0.000
Listens to radio	-0.008	0.992	0.364
Resident since 1993	-0.034	0.966	0.141
Ideal number of children	0.026	1.027	0.000
<i>Ethnic Effects</i>			
Kalenjin	0.105	1.110	0.015
Kamba	0.031	1.032	0.705
Kisii	0.016	1.016	0.811
Luhya	0.124	1.132	0.022
Meru Embu	0.019	1.020	0.829
Miji_swah	-0.018	0.983	0.877
Luo	0.157	1.170	0.015

Table 8: Children Ever Born: Zero Inflated Poisson.

	$\hat{\beta}$	$e^{\hat{\beta}}$	p-value
<i>Endogenous Effects: Cluster-level</i>			
$\bar{C}_{e,cl} - C_i$			
Kalenjin	0.014	1.014	0.086
Kamba	-0.001	0.999	0.942
Kikuyu	-0.037	0.963	0.018
Kisii	0.007	1.007	0.742
Meru-embu	0.019	1.019	0.396
Mijikanda-Swahili	-0.000	1.000	0.993
Luhya	-0.002	0.998	0.780
Luo	0.002	1.002	0.914
<i>Exogenous Effects: Cluster-level</i>			
$\bar{E}_{e,cl}$			
Kalenjin	-0.358	0.700	0.145
Kamba	-0.459	0.632	0.070
Kikuyu	-0.396	0.673	0.282
Kisii	-0.808	0.446	0.121
Meru-embu	0.031	1.032	0.934
Mijikanda-Swahili	0.166	1.180	0.860
Luhya	0.074	1.077	0.594
Luo	0.383	1.467	0.251
<i>Cluster-level Controls</i>			
Media accesss (radio)	-0.014	0.987	0.611
Discussed FP with friends etc.	-0.013	0.987	0.638
Elec. in HH	-0.026	0.975	0.183
Water (nat. source)	-0.048	0.953	0.495
Water (piped into HH)	-0.010	0.990	0.885
Water (public source)	0.096	1.010	0.899
<i>Endogenous Effects: Ethnicity</i>			
$\bar{C}_e - C_i$			
Kalenjin	0.164	1.179	0.000
Kamba	0.202	1.224	0.000
Kikuyu	0.256	1.292	0.000
Kisii	0.194	1.214	0.000
Luhya	0.188	1.206	0.000
Meru-embu	0.200	1.222	0.000
Mijikanda-Swahili	0.202	1.224	0.000
Luo	0.173	1.189	0.000

Table 8: (continued)

	$\hat{\beta}$	$e^{\hat{\beta}}$	p-value
<i>Household Level</i>			
Number of other wives	-0.003	0.997	0.822
Total HH size	0.010	1.010	0.655
Iron roof	-0.029	0.971	0.000
Tiled roof	-0.120	0.887	0.106
Other roof	0.018	1.018	0.801
Wood floor	0.000	1.000	0.997
Tiled floor	0.024	1.024	0.702
Cement floor	-0.012	0.990	0.225
Toilet facility	-0.001	0.999	0.325
Has radio	-0.000	1.000	0.968
Has TV	0.022	1.022	0.084
Has telephone	0.023	1.023	0.223
Has bicycle	-0.004	0.995	0.531
Rooms (sleep)	-0.005	0.995	0.107
<i>Individual Level</i>			
Age	0.178	1.195	0.000
Current Age	-0.002	0.998	0.000
Education (primary)	0.054	1.056	0.000
Education (secondary)	-0.026	0.974	0.035
Education (higher)	-0.066	0.936	0.084
Listens to radio	0.006	1.006	0.114
Resident since 1993	-0.020	0.980	0.072
<i>Inflation Sub-Model</i>			
Current Age	-1.499	0.223	0.000
IdealFszero	-1.910	0.148	0.023
Kalenjin	-0.335	0.715	0.513
Kamba	0.050	1.051	0.922
Kisii	0.979	2.661	0.126
Luhya	-0.010	0.990	0.986
Meru-embu	1.003	2.725	0.118
Miji-Swah	-0.700	0.496	0.265
Luo	-2.577	0.076	0.000
<i>Ethnic Effects</i>			
Kalenjin	0.299	1.349	0.000
Kamba	0.123	1.131	0.001
Kisii	0.117	1.125	0.002
Luhya	0.260	1.297	0.000
Meru Embu	0.095	1.100	0.014