Social Interactions, Ethnicity, Religion and Fertility in Kenya

Sriya Iyer
Faculty of Economics
University of Cambridge

Melvyn Weeks
Faculty of Economics
University of Cambridge

March 5, 2020

Abstract

Reproductive externalities are important for fertility behaviour in Kenya. We identify from anthropology structural forms of social interaction operating across individuals belonging to different ethnic and religious groups on the number of children ever born. We use the 1998 Demographic and Health Survey, combined with primary meteorological data on Kenya, and GMM methods, to show that social interactions effects by ethnicity are important over and above an individual's characteristics such as their religion to explain variations in fertility. Our findings have implications for policy debates in Kenya and in other developing countries about ethnic, religious and other differences in fertility behavior.

---

1 Introduction

Both economists and demographers examine the balance between economic and non-economic factors in orchestrating a fertility transition, as witnessed in historical European populations and in some East Asian economies over relatively short periods of time. Classic studies of fertility such as the Princeton Fertility Project highlight that the transition to low fertility in historical European populations occurred in a variety of socio-economic and institutional contexts, with a significant role being played by the local social environment (Coale and Watkins (1986), Federici, Mason, and Sogner (1993)). More recent studies have emphasised the role of the social climate and the influence of social interactions on demographic behaviour (Manski and Mayshar (2003), Munshi and Myaux (2006)). These empirical findings have led many economists to focus on modelling the influence of social interactions on contemporary fertility transitions (Durlauf and Walker (1999); Manski and Mayshar (2002)) and on the reproductive externalities and coordination failures associated with fertility behaviour per se ((Dasgupta (2000); Kohler (2001); Iyer and Velu (2006)).

The total fertility rate in Kenya fell from an average of 6 births per woman in 1948 (Ajayi and Kekovole (1998)) to approximately 4.56 in 2010 (CIA (2010)). However, the most significant declines in fertility occurred over a twenty year period: 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; Robinson 1992). So it is useful to examine how fertility declined in this society. Moreover, our focus on ethnicity and fertility in Kenya mirrors policy discussions in Kenya that surrounded the release of the controversial Kenya Census figures in 2010 which included ethnicity and tribal affiliations for the first time, following the ethnic tensions of 2007-2008. In light of the sensitivity of these issues, we focus our study of ethnicity, religion and fertility, and our data, on an earlier period in Kenya's demographic history.

Our focus on religion and fertility is motivated by a wider demographic literature that has argued that this issue needs further investigation in developing country contexts (Iyer 2002; Iyer (2019)). In a broad context fertility is higher across the world for Muslims than for Christians (Iyer 2019).This paper presents a context where both Muslims and Christians co-exist, but there are also ethnic differences which may swamp the religious differences. Kenya is a country which displays a great deal of heterogeneity by ethnicity but also by religion. In fact, it is a very interesting example of a society in which religion and ethnicity have interacted to influence behaviours. While ethnic groups have predated the influence of religions like Christianity and Islam in Kenya, there are also indigenous religions here which are less common now, but which predate Christianity and Islam. Hence it is quite useful also to examine how historically well-defined social groups, whether delianated by religion or ethnicity, or both, interact with each other to influence outcomes such as demographic behavior.

Economists consider the impact of social interactions within predetermined groups (Akerlof (1997); Brock and Durlauf (2000), Brock and Durlauf (2001)), and how group formation results from social interactions with particular emphasis on residential neighbourhoods,

geographic proximity and dynamic group formation (Borjas (1995); Ginther, Haverman, and Wolfe (2000); Conley and Topa (2002)). Our study examines social interactions in predetermined groups in Kenya such as religious groups and ethnic groups, that is motivated by the anthropological literature on Kenya. Important previous studies which have examined fertility rates and education in Africa have highlighted the significance of social interactions in Kenya (Behrman, Kohler, and Cotts-Watkins (2002), Kravdal (2002), Rutenberg and Cotts-Watkins (1997))

In line with these studies, 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 also take these studies forward by suggesting that the social environment impinges on a woman's behaviour by multiple interactions both at the household level and at the level of other reference groups such as religion and ethnicity. We highlight a crucial point often overlooked in the theoretical and empirical literature on this issue: namely that there are *different levels at which social interactions occur.* In doing so we consider the possibility that both local and global forms of social interaction may coexist (Horst and Scheinkman (2003a)). We quantify the effects of social interactions and group membership on the number of children ever born, highlighting whether these effects are mediated through household composition, ethnic affiliation, religion, a neighbourhood cluster effect, timing effects, or some combination thereof.

In Section 2 we provide an overview of ethnic and religious groups in Kenya, highlighting the characteristics which are relevant for fertility behaviour. Section 3 summarizes social interactions and fertility behaviour including theory, measurement and estimation. Section 4 presents our demographic data. Section 5 discusses how we use unique district-level rainfall data from Kenya's meteorological rainfall stations over time to identify variations in fertility by ethnic group. We discuss our results in Section 6, and Section 7 concludes.

2 Anthropological overview of groups in Kenya

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

Population   The population of Kenya consists of three groups - the Africans, and to a lesser extent, the 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 original 1969 *Population Census* there are 42 different ethnic African groups in Kenya (Rep (1970)). According to the most recent 2009 Census, the five largest groups, accounting for over 75% of the population, are the Kikuyu (17%), Kamba (10%), Kalenjin (13%), Luhya (14%) and Luo (10%(CIA (2010)). Their population, according to recent controversial Kenyan census is the Kikuyu at 6.6 million, the Luhya at 5.3 million, the Kalenjin at 4.9 million, the Luo at 4 million, and the Kamba at 3.8 million (*Kenya National Bureau of Statistics*). There are also smaller groups such as the Mijikenda (5%), Meru (4%), Turkana (2.5%), Maasai (2.1%). About 9% of the population are smaller indigenous group below 1% each, and the Arabs, Indians and Europeans are about 1%.

The Kenyan population is 41 million in total in 2011, growing at an average growth rate of more than 3% per year from 1980-2010 and at 2.7%CIA (2010). These Census data from Kenya thus depict a steady increase in the total population of Kenya from 1948 to the present. The total fertility rate in Kenya has fallen from an average of 6 births per woman in 1948 (Ajayi and Kekovole (1998)) to approximately 4.56 in 2010 (CIA (2010)). 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; Robinson 1992). 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. 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 (although this could also be driven by lower costs of children) and it is reported use contraception much less than other groups (Watkins, Rutenberg, and Green (1995)).

Religion   The predominant religion in Kenya is Christianity (Pew (2019)). In 2019, an estimated 84.8% of the total population of Kenya is Christian. Islam is the second largest religion in Kenya, and practiced by 11.1% of the population. Other religious faiths observed in Kenya today are the Baha'i faith, Buddhism, Hinduism and various traditional religions.

Roman Catholicism first came to Kenya in the fifteenth century with the advent of the Portuguese, and then spread rapidly during the 20th century due to missionary activity. Today, the main Christian denominations in Kenya are Protestant and they comprise about 47.4% of the country's religious population. The Roman Catholic Church is about 23.3% of the population. The non-Protestant and non-Catholic groups constitute approximately 11.8% of the population. Most Muslims in Kenya are Sunn Muslims. Traditional religions include the Kikuyu religion and the Mijikenda religion and there are also the Maasai, Turkana, Samburu and Pokot tribes who adhere to traditional religions. The numbers of Hindus in Kenya is small and only about 0.14% of the population. In our DHS sample, there are three main religions which comprise 96.5% of household. The dominant religion is Protestant (66.2%), followed by Catholic (27.8%) and Muslim (2.5%). 3.2% of households report no religious faith. 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. In Tables 8, 9, and 10 presented later in this paper, we examine the influence of social interactions with respect to ethnicity, also controlling for the effects of religion. As our results show later in the paper, despite these significant religious variations, religion does not matter as much as ethnicity in these contexts. And we speculate that this may largely be because the ethnic identities predate the advent of Christianity and Islam in this society.

Region   Ethnic boundaries in Kenya, to a very large extent, are coterminous with political and administrative boundaries (Fapohunda and Poukouta (1997)). Today the regional breakdown of the different ethnic groups is as follows: the Kalenjin reside in the Rift Valley; the Kikuyu live in the Central region, but have also migrated to Nairobi and the Rift Valley. The Meru/Embu reside in the North and East. The Luhya live in the Western province, but have also migrated to Nairobi and Mombasa. The Luo live in Nyanza, with Kisumu as their capital, but have also migrated to the Rift Valley. The Kamba live in close proximity to Nairobi and exhibit ethnic affiliation to the Kikuyu. The Mijikenda/Swahili live in the Coast province. The Meru and Embu groups neighbour the Kikuyu to the north and east. An early study suggests that 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.[1]

---

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

Residential pattern of settlement   One of the most notable features of these ethnic groups is their pattern of residence.[2] 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 division[3] (Berg-Schlosser (1984): 50). Supplementing the organisation of Kikuyu society based on kinship, there is also a geographical demarcation. The *'itura'* or village consists of groups of families. This is further subdivided into those living on the same ridge. The pattern of Luo settlement is similar to other communities, although their homesteads usually consist of families who own land, dispersed as a safeguard against climatic conditions. The Kamba are also agriculturists, and their pattern of settlement is similar to the Kikuyu (Middleton and Kershaw (1965)). The traditional land-owning unit is the extended family. In terms of territorial residence, the most basic unit is the *ukambani,* which is a homestead that comprises several extended families. Several of these are grouped to form a *kivalo* within which social interactions, especially marriage, take place. For the Mijikenda, the main form of residence is the *mudzi* or village which consists of groups of agriculturists and fisherman.

Clan organisation   These tribal groups are also broadly characterised by three main features that have relevance for social interactions: the importance of the family group, the clan, and the system of age-grading[4] (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

---

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

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

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

the paternal line, often for up to seven or eight generations. In addition to its functions as the most important traditional land-holding unit in Kikuyu society, the *mbari* is an important social reference-group for many Kikuyu and still plays an effective role in many economic and social relationships including the more modern ones.' (Berg-Schlosser (1984): 53). We would therefore expect that the endogenous social interaction effects among this group would emerge as being particularly strong. The Luhya do not have as powerful a clan organisation as some of the other groups (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).

Polygyny One of the interesting features of African societies including Kenya is the incidence of polygyny. Polygyny can have two impacts on fertility: first, it can increase fertility if there is more help provided with child-care, creating a micro-level externality within the extended family household. Second, if there is greater discussion of family planning issues, then this might work to reduce fertility. Alternatively if high-fertility norms are espoused within the extended household then the effect might be to increase fertility.A number of recent studies have examined the role of polygynous marriage norms in influencing demographic behaviour Hogan and Biratu (2004). There are 30% of polygynous households in the Kenya DHS sample that we consider in this study, so these were households that had more than one wife living in the household. This is not related to the household's religious affiliation. The rationale for looking at polygyny is that if other household members help with child care, or indeed as with other African societies where 'fosterage' is common, other residents have an important bearing on a woman's total fertility.A number of other demographic studies of poor societies (see, for example, Iyer (2002)) have found that the role played by residents within the household or within close proximity, such as friends and neighbours, is important for fertility. We acknowledge that the effects of polygyny might be even more complex than this. For example, unobserved differences among women who enter polygynist unions might be correlated with unobservables in the fertility equation. In general, the frequency of polygynous unions vary a great deal across Africa: DHS data show that they are between 30-50% in West Africa, 20-35% in East Africa, and less than 20% in Southern Africa. So our sample which has 30% polygynous unions, is typical of East Africa where Kenya is situated geographically.

Mortality Infant mortality and mortality caused by HIV/AIDS also would affect fertility rates. We examine infant mortality and the awareness of HIV/AIDS in the population. which are included in order to control for the fact that across all ethnic groups in Kenya, as like other countries in Africa, there have been a large increase in HIV-related deaths which might also affect fertility. For example, recent data show that 7.1-8.5% of adults aged 15-49 in Kenya are currently living with AIDS. Higher AIDS-related deaths may increase fertility if many children are lost to this. There is also the possibiility of reverse causality in that low CEB groups may use contraceptive methods more and hence this is a cause of their low HIV rates as well. It is also important to examine infant mortality by ethnicity

in Kenya. This is because demographers think that higher infant and child mortality is frequently associated with higher fertility rates if parents want to replace children who are lost to child death, or if they perceive that the average mortality rate in their region is high. So it is important to control for mortality changes which might also influence fertility in this country. As regards differences in mortality by ethnic groups in Kenya, Tabutin and Akoto (1992) found that the probability of infant mortality among the Luo was higher than among the Kikuyu. Brockerhoff and Hewett (2000) argue that in Kenya mortality differentials by ethnic group exist and are related to several factors. Migratory patterns influence the spread of disease, and economic and other conditions.For example, they contrast the Kalenjin who are more sedentary with the Luhya who migrate more and are hence more prone to exposure to infectious diseases. They also highlight the Kikuyu who have better access to health and education which meant that they first experienced initially at least the highest declines in mortality of any ethnic group in Kenya. For this reason, the Kikuyu also had very low child mortality (the under-5 mortality rate for Kikuyus was 36.1 per 1000 live births in 1989-1993), but for non-Kikuyu groups, this was 125.1 deaths per 1000 live births. They attribute this to resource concentration, and high altitude that prevents the spread of malaria among Kikuyu children relative to other groups.

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

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

Politics of ethnicity Events in Kenya in 2007-2008 highlighted the importance of the politics of ethnicity in this country. Differences in the economic performance of ethnic

groups can be traced to their evolution in Kenyan history and to economic decisions made by these groups (Ferguson and SrungBoonmee (2003)). Ethnic identity in Kenya has always been very strong historically, and it continues to be a very potent force in Kenya's politics even today. For example, the Kalenjin, who are a heterogenous ethnic group and who live primarily in Kenya's Rift Valley province (Were (1967)) are closest to Nairobi geographically, and they are politically and economically Kenya's dominant ethnic group (Kenyatta (1966)). Their location in the eastern Rift Valley allows them to enjoy a level of political proximity that, in part, determines their (relatively) superior economic status. Agricultural innovation first arose among the Kikuyu who also absorbed land reform more readily than other groups (Meck (1971): 27). Historically, the Kikuyu were one of the first ethnic groups in Kenya to absorb European-style capitalism in the form of wage labour and participation in the monetary economy, so in contrast for example to the Maasai whom they neighbour, the Kikuyu's subsequent economic performance has been much better (Ferguson and SrungBoonmee (2003)). [5] 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 ethnic identity has also led, more worryingly, to ethnic clashes in 1992-93, 1997 and 2007-2008.

3 Social Interactions and Fertility: Theory, Measurement and Estimation

3.1 Theory

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)). The standard Beckerian atomistic utility function (Becker (1981), Willis (1973)) is given by

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

where U_i denotes the utility function of household i, c_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.[6]

---

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

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

ε_i allows for imperfect information on the part of the analyst. This is subject to the usual budget constraint

$$I_i = c_i p_c + \mathbf{z}_i \mathbf{p}_z, \tag{2}$$

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

The extension of the atomistic utility model to incorporate social interactions allows us to capture that component of individual utility attributable to a fertility choice c_i, that is dependent upon the fertility choices of others (see, for example, Becker and Murphy (2000)). The 'others' in question may be those resident in the individual's household or in the local community; for example, in the individual's street or neighbourhood.

It is important to distinguish between *local* and *global* interactions as individual behaviour may be determined by the decisions of a larger ethnic group. In a local interaction, individuals have incentives to conform to the behaviour of a small number of appropriately defined neighbours, whose actions they can 'observe'. In contrast a global interaction is where individuals face incentives to conform to the 'expected' behaviour of a common reference group (Horst and Scheinkman (2003b)), whose behaviour we cannot observe. It is possible that these expectations are channelled through the media, or through other norms of behaviour such as knowledge of the customs and traditions of particular ethnicities. There is thus the possibility of multiple social interactions - namely that social interaction is both localised within a specific location, and more diffuse across a larger, geographically diffuse sub-population.

Following Manski (1993), three hypotheses are advanced to explain why individuals belonging to a common reference group behave similarly.[7] In a fertility context, there is an *endogenous* effect if, ceteris paribus, children ever born (CEB) to a woman varies 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, CEB to a woman varies with the average characteristics of her ethnic group or locality. For example, the existence of educated neighbours may foster positive attitudes towards smaller family size.

Education exercises a prominent effect on fertility behaviour. For example, individual educational attainment affects 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. There is also an iconic value of education as 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 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

by her.

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

10

continue one (Srinivas (1994)).

Finally, there is a correlated effect if, ceteris paribus, women in the same ethnic group or locality have similar CEB because they are, for example, similarly wealthy.

The importance of differentiating between these three hypotheses can be seen by examining their respective policy implications. Consider, for example, an intervention to provide free contraception to some members of an ethnic group or a number of neighbourhoods. If there are endogenous effects, then an effective policy of free contraception may both directly reduce the fertility of the recipients but, as their fertility decreases, indirectly reduces the fertility of all other members of that ethnic group or neighbourhood. Exogenous effects and correlated effects will not, in general, generate this kind of social multiplier effect.

3.2 Measurement

As outlined in Section 2, using a number of anthropological sources we have identified that within rural populations in Kenya, fertility decisions are conducted within families which are members of ethnic groups and/or religious groups, and which reside mainly in homesteads that are located in relatively small clusters. We postulate that these allegiances, form the basis of social interactions. We focus upon the role of two reference groups. First, an *ethnicity* effect derives from the fact that the status of women and attitudes towards children differ across ethnic groups, with differential evaluation of the psychological costs and benefits of bearing children.[8]

Second, in the case of Kenya, situated between the individual and the ethnic group, are small *clusters of households* within which physical proximity dictates that individuals directly observe and bear the costs of the decisions of others. The clusters are homogenous in that they usually consist of the same ethnic (and/or religious) group. 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 interactions.[9] 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.

3.3 Estimation

In motivating a modelling strategy we first consider the observed data. For each ith woman we observe $\{c_i, \mathbf{x}_i, L_{ij}\}$, where c_i represents a count of the total number of children ever born, $\mathbf{x}_i$ is a vector which includes individual characteristics, together with ethnic, cluster, and household attributes; $L_{ij}, j = 1, ..., J$ denotes the jth location in which i resides.

---

[8]There has been a great deal of research on the effects on fertility of religion and ethnic group membership (see Iyer (2002) for a detailed discussion of this literature).

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

Our sample, restricted to women between 30 and 49 years of age, is designed to capture the decision-making of women who have *completed* their fertility, thereby excluding relatively young women for whom timing effects may be confounded with changes in fertility. The presence of timing effects, for example younger cohorts marrying at later ages, are important in that they determine, in part, the difference between the observed number of children for a given woman and the mean number of children in the appropriate reference group. This is central to our thesis, namely that changes in this difference will impact upon CEB due to an interaction effect.

Dynamic models of fertility consider birth spacing, birth timing, the influence of uncertainty, and essentially view fertility decisions sequentially (for a more recent example see Iyer and Velu 2006). Seminal dynamic models of fertility such as Wolpin (1984), Newman (1988) and Rosenzweig and Wolpin (1985) address questions such as the importance of birth timing and spacing. The Wolpin model examines birth spacing, birth timing, and intertemporal tradeoffs. In this model an individual maximises the expected value of the discounted sum of period utility over their lifetime by choosing a sequence of births and consumption, subject to infant mortality and shocks to household income.

3.3.1 Model Specification

In this section we consider the specification of the conditional mean, and in particular the specification of the exogenous and endogenous components of fertility behaviour.

Given that count data is both integer valued and heteroscedastic by construction, the Poisson density represents a natural benchmark to model count data. However, the equidispersion property, namely that the conditional mean is equal to the conditional variance, is generally restrictive. In this paper we utilise the Generalised Method of Moments (GMM) estimator. A GMM estimator[10] is particularly suited given that it allows for unobserved heterogeneity without a full set of parametric assumptions, and readily accommodates endogenous regressors.

We write the conditional mean as

$$E(c_i|\mathbf{x}_i, \omega_i) = \exp(\, \mathrm{cf}_{i,g}\delta + \mathrm{ed}_g\gamma + \mathbf{h}_i'\theta + \mathbf{p}_g'\kappa + \mathbf{q}_i'\eta + \omega_i \,). \tag{3}$$

c_i denotes the number of children (CEB) ever born to the ith woman. (3) includes the standard set of individual ($\mathbf{q}_i$), household ($\mathbf{h}_i$), and cluster ($\mathbf{p}_g$) characteristics, with associated vectors of parameters η, θ, and κ. cf_{ig} represents conformist interaction effects attributable to the fertility choices of other women in the same group. ed_g denotes exogenous interaction effects due to education which we measure by the mean level of education in the cluster resided by the individual. δ and γ are the associated scalar parameters.

Exogenous Interaction Effects

We represent exogenous effects attributable to the education of other women in the same cluster using $\mathrm{ed}_g = N_g^{-1} \sum_{i \in g} \mathrm{ed}_i$ with ed_i denoting the education level of the ith woman residing in the same cluster. It is argued that the effect of education is iconic, with some

---

[10]See Appendix I for for further details.

individuals aspiring to the attributes of other higher educated members within their cluster. We choose to model these education interactions at the level of the cluster as higher educated women in the locality may be leaders and opinion-makers who influence others who are less educated to adopt low fertility norms. Here we consider the proportion of women in the cluster who had completed higher education as representative of this exogenous education effect on fertility.

Endogenous Interaction Effects

We extend the standard atomistic utility model of fertility to capture that component of individual utility attributable to a fertility choice that is dependent upon the fertility choices of others. Individuals may have incentives to conform to the behaviour of a small number of appropriately defined neighbours whose attributes they can 'observe'.

In the conditional mean specification (3) endogenous conformist effects are represented by $\text{cf}_g = E_i(c_g - c_i)$, where c_i denotes the number of children (CEB) ever born to the ith woman, and $E_i(c_g)$ is the subjective expectation of mean CEB held by the ith woman resident in cluster g.[11] Restricting the sample to those women aged 30-449 and thereby confining our analysis to completed lifetimes, circumvents the problem of pooling data over age cohorts. For older women we can appeal to rational expectations and set the subjective expectation of cohort fertility equal to the mean observed level of fertility, $E_i(c_g) = \bar{c}_g$. However, this becomes problematic when considering subjective expectations of younger women who are not near the end of their reproductive careers.

In Appendix III we outline our identification strategy which explains how we isolate the effect of social interactions from other determinants of fertility behaviour. We do this by first outlining the identification problem in the linear-in-means social interaction model as originally considered by Manski (1993), and more recently by Blume, Brock, Durlauf, and Ioannides (2010) and Ioannides (2008). Since the question of identification is well understood in the linear framework, this will provide a useful point of departure.

3.3.2   Household-level and Individual-level Controls

We control for a number of household and individual characteristics that we believe impact fertility decisions to ensure that the endogenous and exogenous effects are robust to variations in other characteristics.

Individual-Level Controls

Individual-level controls include the age of the woman measured in years and an age-squared variable accounting for the non-linearity associated with age-related variables. The education of the woman is included and measured as the highest level of education attained, differentiating between primary, secondary, and higher education effects. The influence of the media was measured as whether the woman listened to the radio at least once a week. This variable was included on the assumption that greater information about contraceptive technology would be available to the woman if she listened to the radio. Messages about this are routinely broadcast on the national radio channel, the 'Voice of Kenya' in English,

---

[11]See Appendix II for further discussion of conformist effects.

Swahili and the vernacular. We also include a variable here for whether the individual knows a place to test for HIV/AIDS. This is to control for their awareness of AIDS, which is an important contributor to mortality in Kenya and which therefore is likely to affect fertility, as also discussed in Section 2.

We also include the following measures: age at first marriage, and knowledge of any method of contraception. We expect that this will capture the extent to which the age at first marriage is greater for younger relative to older women. With the success of recent family planning programmes in Africa, knowledge of contraception is also likely to be greater for younger women than older ones. We also control explicitly for religion of the respondent.

Household-Level Controls

One notable problem with 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, we constructed a number of indicators which control for the economic status of households. The quality of roof construction was measured by whether the household had a tiled roof or not. The quality of the floor was measured by whether the household had a tiled or wood floor or not. The status of the household was also proxied by ownership of consumer durables including whether or not the household owns a radio, or a television.

Two other important household-level attributes that we utilised are (i) a measure of infant mortality, and (ii) reported awareness of HIV/AIDS in the population. These factors are included in order to control for the fact that Kenya has witnessed a large increase in HIV-related deaths over the time period under study which might also affect fertility. And that different religious groups might have different norms about contraception which can again affect both CEB and CEB rates. The DHS data provide data on infant mortality but because they do not have exact numbers for the HIV/AIDS-related deaths, we use reported awareness of HIV/AIDS in the population for our measure of how pervasive this is likely to be by region.

As discussed earlier, polygyny can have two impacts on fertility: first, it can increase fertility if within polygynyst households provide more child-care, creating a micro-level externality within the extended family household. Second, if there is greater discussion of family planning issues, then this might work to reduce fertility. Alternatively if high-fertility norms are espoused within the extended household then the effect might be to increase fertility, again which might also be related to religion. There are 30% of households in the Kenya sample which are polygynous, so these were households that had more than one wife living in the household. We acknowledge that the effects of polygyny might be even more complex than this. For example, the coefficients on all other estimated parameters may vary by family structure. And unobserved differences among women who enter polygynist unions might be correlated with unobservables in the fertility equation. But we prefer to adopt a more parsimonious specification controlling for these effects. In order to do so, we also control for the number of usual residents in the household as a proxy for polygynous households.

3.3.3 Cluster Attributes

We control for a number of attributes at the level of the cluster. Critical in this respect is access to water and fuel infrastructure. Based upon the work of Dasgupta (2000), better access to water and fuel reduces the demand for child labour to collect them, and that this is turn reduces the demand for children, and hence the fertility rate (Iyer (2002)). Access to fuel was measured by a mean-level effect for access to electricity. Access to water infrastructure at the cluster-level was measured according to whether access to water was piped into the residence or not. In addition, an additional cluster-level attribute included was whether or not a radio was listened to at least once a week. This variable was estimated at the cluster level by aggregating up whether or not an individual listened to a radio. Finally since the process of interaction will depend upon the size of the group, we control for the population size of the cluster

4 Data and characteristics of DHS survey

Our sample, taken from the Kenyan Demographic and Health Survey (KDHS)Ken (1999), covers 88% of the population. The survey contains interview data for 7800 women aged 15-49. It contains women from all of Kenya's large ethnic and religious 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. The representation of each group in the sample is similar to its representation in the whole population. The DHS data includes information on ethnic groups, of which the six main groups are the Kikuyu, Kalenjin, Luhya, Luo, Kamba and Kisii. The Kikuyu is the largest, with 17.9% of the population.[12] In the data used in this study we consider individual-level data only for the 5994 women residing in a rural cluster, and who did respond. 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.[13]

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.[14] 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

---

[12]The DHS sample 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.

[13]No household from Nairobi were present in our estimation sample. We therefore exclude Nairobi from the summary tables.

[14]6 of the clusters could not be included in the survey due to inaccessibility (DHS and Macro International 1999: 180).

cluster was recorded between November 1997 and February 1998. From the remaining 530 clusters, a systematic sample was drawn of, on average, 22 households in urban clusters and 17 households in rural clusters. This formed a total of 9465 households. In these households, all women age 15-49 were targeted for interview. Response rates varied by province from approximately 88% to 99% (for more details see KDHS, 1998 p.180).

Table 1 presents the mean CEB by region and ethnicity. Average CEB varies from a low of 2.58 in Central to a high of 3.9 in Nyanza. For a number of regions, including Nyanza and Western we observe notable intra-regional variation in CEB by ethnicity.

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. As shown in Table 1, the differences in fertility by ethnic group are clearly very large.

Our sample also records the religion of the household. There are three main religions which comprise 96.5% of household. The dominant religion is Protestant (66.2%), followed by Catholic (27.8%) and Muslim (2.5%). 3.2% of households report no religious faith.[15] Table 2 presents the mean CEB by region and religion. The data reveal variation in mean CEB induced by the combination of religious and regional effects.

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 38% of women who use contraception and the most widely used methods are contraceptive injectables, the pill, female sterilization and periodic abstinence. Contraceptive use does however vary greatly by region: while there are 57% of women in the Central province who use contraception, only 24% 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.

We also observe some variation by religion. 38% of Christian and 39% and Catholic women use contraception, with the number falling to 27% for Muslim women and 39% forthose who had no religion.

Table 3 presents mean years of education by region and ethnicity. The mean years of education are highest for the Kikuyu at 8.1 years while the lowest at 4.2 years is among the Mijikenda/Swahili groups. Table 4 presents the distribution of total children ever born in the sample. The main range of total children ever born is between 0 and 8 children in this sample.

---

[15]For a relatively small number of households the religious affiliation is not reported.

5 Metereological Data as an Instrument for Fertility

In this section we present a justification for the use of the rainfall instruments. In Section 5.1 discuss how the instruments were constructed.

We contend that for tropical countries like Kenya which experience hot weather and exhibit less fertility control than in developed countries, properties of the distribution of rainfall, and in particular the mean and variance, represent valid instruments for the observed variation in fertility.

The use of rainfall data as an instrument is based on a number of arguments. First, in low-income countries in which fertility regulation is not practiced extensively, and in which parents calculate the costs and benefits of having an additional child based on income (Becker (1981)), rainfall variations will influence fertility behaviour. For example, in periods of drought when income is likely to be low, parents might postpone childbearing (Lam and Miron (1996)).

Second, high temperatures will have an effect on conception and thereby on fertility. In the main, since the Kenyan ethnic groups tend to reside in distinct regions then geographic variation in rainfall is likely to explain variation in fertility rates by ethnicity.

In addition to the impact of mean rainfall, we argue that the seasonality of rainfall exerts an additional independent effect on fertility decisions, and as such represents an additional instrument. The argument for the use of this instrument is based on the notion that seasonality of rainfall proxies uncertainty. Recent papers in economic demography have also been examining the impact of uncertainty and the real options approach in order to understand decision-making about fertility (see, for example, Iyer and Velu (2006); Bhaumik and Nugent (2010) The impact of uncertainty on fertility can be either positive or negative. If parents view children as an 'investment' good then a decrease in uncertainty can reduce the insurance value of having children and thus decrease the net benefit of having children (Iyer and Velu (2006)). Alternately, if children or 'child services' are viewed purely as a 'consumption' good, then a decrease in uncertainty is likely to result in an increase in fertility as the desired number of children rises.

Combining these arguments, we note that in areas of abundant rainfall, which are on average cooler and have lower temperatures, we expect *a priori* that these areas will also be areas of higher fertility. This is because in areas of abundant rainfall, income will be high. Areas that are very arid and receive on average very poor rainfall are also likely to be those with very high hot weather temperatures, and will exhibit lower fertility as conception rates are likely to be low. We would therefore expect to observe higher fertility in better rainfall-fed districts of Kenya as compared to those which have lower rainfall and greater seasonality of rainfall as well.

5.1 Constructing the rainfall instruments

In this study we use unique historical meteorological data on monthly rainfall collected for us by the Kenyan Meteorological Department from thirty-one meteorological stations in Kenya at 5-year intervals, from 1930 to 2000. This data was aggregated into yearly

average rainfall and its standard deviation by rainfall station in each Kenyan district. In considering how we construct our instruments it is important to consider the nature of the model. Given that we are using a static model of completed fertility, then in addressing the problem of endogeneity it makes sense to instrument the conformist endogenous effects with accumulated lifetime rainfall. In terms of the impact of rainfall on fertility behaviour, a given household is likely to make fertility decisions (i.e. the number and timing of children) which reflect both the geographic and temporal distribution of rainfall. However, in specifying a one period static model of completed fertility, requires that we align total fertility with total rainfall faced over the lifetime and construct instruments accordingly.

Below we outline the matching exercise that allowed us to assign to each woman the total rainfall faced over her reproductive lifetime i.e. during her childbearing years. The matching of the metereological and the DHS data proceeded by assigning, for each observation in a particular district, rainfall data which minimised the Euclidean distance between the district capital and the nearest metereological station. [16]

To quantify the impact of uncertainty on fertility, we construct an index for the seasonality of rainfall across the meteorological stations in Kenya over the course of a year. For each year the index takes the difference between the average rainfall in the wettest and the driest three months, and scales them by the average rainfall for the year for that particular meteorological station (see Foeken (1994)). This is reported as the index of District Rainfall Seasonality (DRS).

Table 5 depicts the geographical variations in rainfall across the different provinces of Kenya. Rainfall is highest on the western coast of Kenya (the Western province and Nyanza) in which fertility rates are also the highest (3.2 and 3.3). As shown in Table 6, rainfall in Kenya is highly seasonal and varies considerably by geographical region. We examine correlations between the total number of children ever born and the average rainfall by district in Kenya. This correlation is positive and is 0.437. For women who have 10 births or fewer, this correlation was even stronger at 0.4598. Table 7 shows for the year 1998 children ever born by province, mean rainfall and rainfall seasonality for the different Kenyan regions.

Looking at Tables 5, 6 and 7 collectively, suggests that lower rainfall seasonality might be associated with higher fertility: the greatest seasonality of rainfall is in the Central province and the eastern province, the regions that also have the lowest fertility. Of course these findings could also be driven by other characteristics of these provinces such as income or education, and in the models presented below we examine the effect of our rainfall instrument after controlling for some of these other factors. However, if we look simply at the correlations between these rainfall patterns and fertility rates across the different provinces of Kenya, it would appear that the areas with the *highest rainfall, and lowest rainfall seasonality, are also those in Kenya which show the highest fertility.*

---

[16]We used latitude and longitude data for the district capitals and those of the meteorological stations.

6 Results

We present our main results in Tables 8 9 and 10 for ethnic, religion and interaction effects respectively. We do not report results for the individual, household or cluster level effects in detail in these tables but these are included in all estimations presented, and also discussed below.

For each variable we report three outputs. Parameter estimates, say β_j, may be interpreted as the proportional change in the number of children ever born due to a unit change in the jth regressor.[17] Underneath and in parenthesis we report the p-value, followed by the $e^{\widehat{\beta_j}}$, the factor change (the incidence-rate ratio (IRR) in the expected count for a unit increase in a given x_j. As an example, a IRR = 1.10 indicates a 10% increase in the expected CEB for a one unit change in x_j. We emphasise this statistic when interpreting our results. Standard errors are robust to arbitrary intra-cluster correlation. We also report the J-statistic affording a test of the exogeneity of our instruments. In Table 8 we report the results for our baseline model with ethnic effects. Table 9 reports our results when we include a set of a religion dummies, replacing ethnic effects. Table 10 reports our results, when we include both religious and ethnic effects.

Ethnicity and Religion Effects

Table 8 reports our results for the baseline model. Each table presents results for three specifications. Specification I incorporates traditional fertility determinants at the individual, household and ethnicity level. Specification II extends I with the addition of ethnic-level exogenous effects (due to education), with specification III adding ethnic-level conformist effects.

Ethnicity effects were significant for all ethnic groups such as the Kalenjin (1.221), the Luhya (1.272), the Kisii (1.154) and for the Luo (1.321), revealing larger fertility effects relative to the Kikuyu which are the reference group. The relative magnitudes of the ethnic effects reflect what the descriptive statistics also show in Table 1 namely that the Kikuyu have the lowest fertility of any of the ethnic groups. Given that in many cases ethnic groups settle in specific regions, it was not possible to separately identify all regional effects.

In Table 9 when we replace ethnicity with religion, the only group for which the effect is significant is for the group with no religion (1.120). In Table 10 when we include both ethnicity and religion, the effects are significant for the ethnici groups Kalenjin (1.216), Luhya (1.270) and Luo (1.323) only, and not significant for any of the religious groups.

Ethnic-Level Education Effects

We consider exogenous education effects as attributable to the effect of observing the level of education of other women in the same cluster. We measure this effect using the proportion

---

[17]This follows since

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

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

of women who had completed higher education, allowing parameters to differ by ethnicity and the age of the woman.[18]

With the exception of the Luhya ethnic group, ethnic-Level education effects in Table 8 are insignificant after controlling for individual educational attainment. They are also unimportant in Table 9 when we substitute religion for ethnicity, and they are only significant for the Luhya group (0.386) when we include both ethnicity and religion in Table 10.

Ethnic-Level Conformist effects

In Table 8 the localised endogenous effects emanating from within the village cluster, are captured by a parameter which, for each ethnic group reflects the impact of the difference between a cluster mean-level fertility and individual CEB - namely the difference between a mean-level and individual CEB for women who live in the same cluster. We observe that for all 6 groups these parameters are significant and positive with effects being strongest for the Kikuyu, ranging from 1.267 for the Kikuyu, 1.247 for the Kamba, 1.209 for the Kalenjin, 1.196 for the Luhya, 1.172 for the Kisii, to 1.131 for the Luo.

We emphasise that these results are consistent with our anthropological understanding of the ethnic groups (as discussed in Section 2). Namely, for those clusters in which individuals were related both by blood and marriage, as for example, among the Kikuyu *mbari*, the endogenous effects on fertility are the strongest of the ethnic group effects. This is important for demographic analyses per se since our results indicate the importance for fertility of including a measure which captures the effects of interactions within ethnic groups, and that they are important over and above individual and household characteristics. This is consistent with anthropological and other studies of fertility outlined in Section 2 which also emphasise the role of these social norms. It is also striking that the Kikuyu who demonstrate the lowest average fertility are also those for whom the endogenous effects are the strongest.

In Table 9 we report a model which is identical except we allow for religious effects but exclude ethnicity effects. Using the Christian religion as a reference category, we find that none of the religious effects are significant. The only group for which conformist effects matter are the Kikuyu (1.135). We note that this model fails Hansen's test of overidentifying restrictions

In Table 10 we report results for a model specification where we allow for religious effects in addition to the role of ethnicity. We again find that religion does not exert any significant incremental effect on the number of children ever born. Comparing the results in Table 10 and 8, we find that the results do not change to any significant degree. For example, in Table 9 with religion, the endogenous effects are all insignificant except for the Kikuyu, while in Table 10 with religion controlled for and ethnicity effects, the endogenous effects for ethnicity are all highly significant. This suggets that the ethnicity social interaction effects

---

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

which we postulate as being important in this society are upheld even after controlling for the effect of religion.

In Kenya there is a fair amount of heterogeneity by religion. However, the question is how this potential effect of religion plays out alongside the generalized norms at the level of the ethnic group, which are much stronger in this society. We speculate that one reason why there may not be an effect of religion is that the tribal groups and tribal norms pre-date the introduction of Christianity and Islam in this society, and may also be associated with historical tribal religions, so this might be one reason why the ethnic effects predominate the religion effects. Hence, we maintain the emphasis on ethnicity and the different channels of effects, even after controlling for religion.

Individual-level effects

We do not report the details of the individual and household-level effects in Tables 8, 9 and 10 but discuss them here. As expected many of the individual-level variables were highly significant: these included the woman's age with an older woman exhibiting higher fertility compared to her younger counterpart. For women, primary and secondary education were significant and positive. Higher education was significant in specification I and as expected, strongly negative.[19] Primary and secondary education has the expected positive relationship with CEB. Also in specifications I and II, if the woman knows a place to test for AIDS, she was likely to have much lower fertility than if she did not. This effect however disappears when the endogenous effects are included. We control for the effect of an individual's religion, but do not see any effects for this, except for those with no religion in Table 9, and these effects disappear entirely when the ethnicity effects are included in Table 10.

Timing Effects

In order to examine the fertility behavior of a group of women whose fertility is largely complete we restrict the sample to women aged 30-49. We also control for age at first marriage, and knowledge of any method of contraception. Age at marriage exerts a strongly negative effect on fertility. Knowledge of contraception was not significant.

Household-level effects

Household-level controls are included so as to reduce the possibility that any inference on household-level interaction effects is not confounded by household-level omitted variables At the household level we make a distinction between interactions within the household which may have implications for fertility, and household-level controls. None of the household-level interaction effects were significant, except the infant mortality variable at the household level in specifications I and II. The significance of the infant mortality variable which measured whether a son or a daughter had died, is consistent with the argument that higher infant mortality usually results in higher fertility as well. Interestingly, when the

---

[19]The base category was women with no education.

endogenous effects are included, the significance of this variable disappeared. This might suggest that there are ethnic variations in infant mortality rates, as discussed further below.

In general we would expect that income would be negatively correlated with fertility. As noted above, one of the problems with the DHS survey is that no direct questions on income were asked of survey respondents. To address this deficiency a number of household level variables were included as controls for income and other characteristics. One measure of roofing quality was included to suggest that if the woman lived in a household that was wealthier, as measured by a better quality roof such as a tiled roof, then her CEB was more likely to be lower, relatively to the base category - a thatched roof. If a woman lived in a house which had a tiled floor, the expected number of children ever born was likely to be lower than if she was resident in a house with a sand floor. The tiled floor variable emerged as very significant. For the roofing quality variables, if the woman lived in a house which had a tiled roof (relative to the base of thatch), then this might have been important but this was not as significant as floor quality. Two variables included in the models which could be construed both as income indicators and as controls for media access are radio and television ownership. Television ownership was strongly negatively significant in specifications I and II, but not radio ownership, in the models with no endogenous effects.

7 Conclusion

Strategic complementarities in fertility decisions implies that a couple's fertility decisions may be dependent on the actions of others in the vicinity or in the society more widely. The mechanisms through which these complementarities occur are through social interactions. This paper has examined social interactions in the context of fertility behaviour in Kenya. We have examined the role of dependencies across individuals that reside in a cluster of households and who locate according to ethnicity. More significantly, we emphasise the existence of multiple channels of social interaction. Our results confirm strongly the importance of ethnic-level effects alongside the traditional determinants of fertility at the individual leve such as religionl. After controlling for the timing effects, endogenous effects, household and individual effects, the exogenous effects are insignificant. This might suggest that social interactions matter more at the level of the extended group - in this instance localised interactions through village clusters and more global interactions through ethnic effects. We do not find any significant effects of religion. In the context of a wider debate about multiple identities, this seems to suggest that the effects of ethnicity may outweigh the effects of religion in this society.

Given their significance as our results show, we feel that more research needs to be done on these effects and the broader relationship between demography and cultural networks such as ethnicity and religion. Our results suggest that in considering a continuum of interactions from the most local i.e. the individual to the most global, the ethnic group, kin groups and the structure of the network of womens' interactions may be more important to fertility decision-making than interactions that take place at the level of the household. This could well be a function of how attitudes and norms related to high fertility might

be changing. For example, as women are better informed about matters related to fertility decision-making beyond their immediate household, whether through the media or through other channels, this knowledge influences norms and exerts greater influence on their decision-making. Our research is of course unable to solve this puzzle precisely because data on networks was not collected in the Kenya DHS survey, but it does suggest that more research is required to identify the precise form of the interactions operating among kin groups, religious and ethnic groups, and how they affect fertility preferences in Kenya and elsewhere, where ethnicity, religion and fertility outcomes are linked.

References

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

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

(2019): "Pew Research Center's Religion & Public Life Project: Kenya," Pew Research Centre 2019.

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

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

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.

BEHRMAN, J. R., H. P. KOHLER, AND S. COTTS-WATKINS (2002): "Ocial Networks and Changes in Contraceptive Use over Time: Evidence from a Longitudinal Study of Rural Kenya," *Demography*, 39(4), 713–738.

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

BHAUMIK, S. K., AND J. B. NUGENT (2010): "Real options and demographic decisions: empirical evidence from East and West Germany," *Applied Economics*, p. 1466.

BLUME, L. F., W. A. BROCK, S. N. DURLAUF, AND Y. M. IOANNIDES (2010): "Identification of Social Interactions," in *Handbook of Social Economics*, ed. by J. Benhabib, M. Jackson, and A. Bisin. Elsevier, North Holland.

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

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

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

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

BROCKERHOFF, M., AND P. HEWETT (2000): "Inequality of child mortality among ethnic groups in Sub-Saharan Africa," .

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

CASE, ANNE, C., AND F. KATZ, LAWRENCE (1991): "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths," National Bureau of Economic Research, Inc, NBER Working Papers: 3705.

CIA (2010): "The World Factbook 2010.," Washington, DC: Central Intelligence Agency, 2010.

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.

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.

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

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

FEDERICI, N., K. O. MASON, AND S. SOGNER (eds.) (1993): *Women's Position and Demographic Change*Oxford. Clarendon Press.

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

FOEKEN, D. (1994): "Climatic Seasonality in Kenya with Special Reference to Coast Province," *Eastern and Southern Africa Geographical Journal*, 5(1), 28–43.

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

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.

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

HORST, U., AND J. A. SCHEINKMAN (2003a): "Equilibria in Systems of Social Interactions," mimeo Humboldt University of Berlin, Department of Mathematics.

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

IOANNIDES, Y. M. (2008): "Empirics of Social Interactions," in *New Palgrave Dictionary of Economics*, ed. by L. E. Blume, and S. N. Durlauf. Palgrave Macmillan, 2nd edition.

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.

——— (2019): "Religion and Demography," in *International Economic Association Series Volume 158*, ed. by J.-P. Carvalho, S. Iyer, and J. Rubin. Palgrave Macmillan.

IYER, S., AND C. VELU (2006): "Real Options and Demographic Decisions," *Journal of Development Economics*, 80(1), 39–58.

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

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

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

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

KRAVDAL, O. (2002): "Education and Fertility in Sub-Saharan Africa: Individual and Community Effects," *Demography*, 39(2), 233–250.

LAM, D. M., AND J. MIRON (1996): "The Effects of Temperature on Human Fertility," *Demography*, 3, 291–305.

MANSKI, C., AND J. MAYSHAR (2003): "Private Incentives and Social Interactions: Fertility Puzzles in Israel," *Journal of the European Economic Association*, 1(1), 181–211.

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

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

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

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

MULLAHY, J. (1997): "Instrumental-Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior," *The Review of Economics and Statistics*, 79(4), 586–593.

MUNSHI, K., AND J. MYAUX (2006): "Social Norms and the Fertility Transition," *Journal of Development Economics*, 80(1), 1–38.

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

RUTENBERG, N., AND S. COTTS-WATKINS (1997): "The Buzz Outside the Clinics: Conversations and Contraception in Nyanza Province, Kenya," *Studies in Family Planning*, 28(4), 290–307.

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

TABUTIN, D., AND E. AKOTO (1992): "Socio-economic and Cultural Differentials in the Mortality of Sub-Saharan Africa," in *Mortality and Society in sub-Saharan Africa*, ed. by E. van de Walle, and M. Pison, G. aSalad-Diakanda. Oxford: Clarendon Press.

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

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.

WINDMEIJER, F., AND J. SANTOS-SILVA (1996): "Estimation of Count Data Models with Endogenous Regressors; an Application to Demand for Health Care," Institut fur Statistik Sonderforschungbereich 386, Paper 20.

Table 1: Children Ever Born By Region and Ethnicity

	Region							
Ethnicity								
	Central	Coast	Eastern	Nyanza	Rift Valley	Western	μ_{CEB}^{E}	N_{SAMPLE}
Kalenjin	8.00	9.00	0.50	6.33	3.34	2.93	3.34	1277
Kamba	3.71	3.41	2.88	4.00	3.80	-	2.93	674
Kikuyu	2.56	3.14	2.50	3.00	3.22	7.50	2.74	958
Kissi	1.67	-	-	2.81	2.65	8.50	2.81	557
Luhya	2.29	3.67	3.00	4.23	3.48	3.36	3.40	913
Luo	2.67	1.00		3.93	3.08	3.30	3.84	741
Meru/Embu	3.60	0.00	2.96	2.00	1.00	-	2.94	444
Miji/Swa	-	3.43	1.50	-	-	1.00	3.41	425
μ_{CEB}^{E}	2.58	3.41	2.90	3.49	3.31	3.36	3.19	
N_{SAMPLE}	728	465	1078	1208	1727	783		5994

μ_{CEB}^{E} (μ_{CEB}^{R}) denote, respectively, mean of CEB by Ethnicity (Region)
− denotes a zero sample count.

Table 2: Children Ever Born By Region and Religion

	Region							
Religion								
	Central	Coast	Eastern	Nyanza	Rift Valley	Western	μ_{CEB}^{Re}	N_{SAMPLE}
Catholic	2.56	3.15	2.78	3.46	3.43	3.35	3.17	1663
Christian	2.56	2.74	2.96	3.50	3.25	3.35	3.16	3996
Muslim	7.00	3.35	3.00	1.50	4.50	4.10	3.47	147
No Religion	2.00	4.27	4.00	7.00	3.56	3.00	4.07	192
μ_{CEB}^{R}	2.58	3.41	2.90	3.49	3.31	3.36	3.19	
N_{SAMPLE}	728	465	1078	1208	1727	783		5994

μ_{CEB}^{Re} (μ_{CEB}^{R}) denote, respectively, mean of CEB by Religion (Region)
− denotes a zero sample count.

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

	Region						
	Central	Coast	Eastern	Nyanza	Rift	Western Valley	μ_{ED}^E
Ethnicity							
Kalenjin	-	7.0	7.3	7.0	6.5	7.0	6.5
Kamba	4.9	6.9	6.7	8.5	6.6	-	6.9
Kikuyu	7.9	9.6	8.7	7.7	7.4	6.7	8.1
Kisii	6.0	5.5	12.0	6.8	6.7	6.0	6.9
Luhya	6.8	8.4	9.6	6.0	6.0	7.1	7.0
Luo	6.7	8.2	8.8	5.9	6.9	8.1	6.3
Meru/Embu	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	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 sample 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: Mean Rainfall (year)

	Western	Eastern	Coast	Rift Valley	Central	Nyanza
1960	159.31	49.80	50.89	63.81	66.95	116.97
1965	155.35	55.09	70.26	54.66	57.04	137.78
1970	179.88	68.03	58.77	73.32	72.37	146.17
1975	167.08	57.20	55.34	101.70	64.09	121.12
1980	133.55	74.96	59.72	79.92	60.58	108.85
1985	188.05	119.69	63.48	90.89	67.07	131.84
1990	167.44	107.35	73.48	88.82	81.10	139.19
1995	167.25	88.19	79.04	89.22	73.98	136.41
2000	117.85	47.57	75.21	71.44	40.79	120.90
2005	125.42	58.58	63.47	75.01	61.00	102.52

Note: Mean rainfall for each province is calculated from the average rainfall of each district within that province.

Table 6: Seasonal Rainfall(year)

	Western	Eastern	Coast	Rift Valley	Central	Nyanza
1960	1.49	2.35	2.17	1.84	1.88	2.34
1965	1.50	2.71	2.21	1.93	2.13	2.21
1970	1.19	2.79	1.84	2.06	1.88	2.54
1975	1.41	2.74	2.51	1.94	2.00	2.20
1980	1.38	2.80	2.36	2.12	2.50	3.00
1985	1.17	2.58	2.22	2.01	2.13	1.98
1990	1.33	2.26	1.85	1.60	1.90	2.41
1995	1.31	2.29	2.57	1.44	1.61	1.47
2000	1.28	2.91	2.52	2.02	2.28	2.25
2005	1.38	2.72	2.10	2.15	2.67	3.02

Note: The index for the seasonality of rainfall across meteorological stations in Kenya over the course of a year. (Foeken, 1994), takes the difference between the averages of rainfall in the wettest and driest three months of the year and scales them by the average rainfall for the year for that particular meteorological station. The resulting calculation is an index per station and for each year in which rainfall is measured. This is reported as the index of District Rainfall Seasonality (DRS). The figures above them average this for each province.

Table 7: CEB Mean Rainfall; and Rainfall Seasonality by Region

Region	CEB (1998)	Mean Rainfall (2000)	Rainfall Seasonality
Central	2.51	40.79	2.28
Coast	2.86	75.21	2.52
Eastern	2.83	47.57	2.91
Nyanza	3.31	120.90	1.59
Rift Valley	3.22	71.44	2.02
Western	3.27	117.85	1.28
ALL	2.98	73.98	2.12

Table 8: Baseline Model: Ethnic and Interaction Effects

	I	II2	III
Dependent variable: CEB			
Ethnicity			
Kalenjin	0.060	0.041	0.200**
	(0.282)	(0.470)	(0.002)
	1.062	1.042	1.221
Kamba	-0.066	-0.083	0.072
	(0.334)	(0.231)	(0.294)
	0.936	0.921	1.075
Kisii	0.018	0.005	0.144
	(0.853)	(0.957)	(0.076)
	1.019	1.005	1.154
Luhya	0.013	0.004	0.240**
	(0.840)	(0.955)	(0.002)
	1.013	1.004	1.272
Luo	-0.052	-0.069	0.279**
	(0.605)	(0.500)	(0.009)
	0.949	0.933	1.321
Ethnic-Level Education			
Kalenjin		0.192	0.347
		(0.711)	(0.684)
		1.212	1.415
Kamba		0.383	0.069
		(0.695)	(0.960)
		1.466	1.071
Kikuyu		-0.989	-1.190
		(0.235)	(0.245)
		0.372	0.304
Kissi		-0.037	-0.570
		(0.955)	(0.189)
		0.964	0.566
Luhya		-0.169	-0.947*
		(0.701)	(0.042)
		0.845	0.388
Luo		0.431	-0.848
		(0.678)	(0.579)
		1.539	0.428
Ethnic-Level Conformist			
Kalenjin			0.190**
			(0.003)
			1.210
Kamba			0.221**
			(0.003)
			1.247
Kikuyu			0.237***
			(0.000)
			1.267
Kissi			0.159***
			(0.000)
			1.172
Luhya			0.179*
			(0.011)
			1.196
Luo			0.123*
			(0.049)
			1.131

* (**) [***] if $p < 0.05$ (0.01) [0.001]

Diagnostics

	I	II2	III	
RMSE	0.346	0.347	0.473	
J-Statistic			7.803	(0.253)

All specifications include the following individual, household, regional and cluster controls
Individual: age, age^2, dummies for primary, secondary, and higher education.

Individual: dummies for listen to radio, know a place to test for aids
Household: dummies for tv and radio ownership, mortality

Household: dummies for: flooring (tile, wood); roof (tiled)
Cluster Controls: media access, electricity, water and size of cluster

Timing: Age at first marriage, knowledge of contraception

Table 9: Religion Model: Religion, and Interaction Effects

	I	II	III
Dependent variable: CEB			
Religion			
Catholic	-0.004	-0.005	-0.001
	(0.852)	(0.801)	(0.952)
	0.996	0.995	0.999
Muslim	0.037	0.045	0.053
	(0.730)	(0.688)	(0.606)
	1.038	1.046	1.055
No Religion	0.101	0.102	0.114*
	(0.057)	(0.059)	(0.015)
	1.107	1.108	1.120
Other Rel	-0.047	-0.044	-0.038
	(0.220)	(0.247)	(0.519)
	0.954	0.957	0.963
Ethnic-Level Education			
Kalenjin		0.384	0.044
		(0.462)	(0.947)
		1.467	1.045
Kamba		-0.436	-0.812
		(0.639)	(0.401)
		0.646	0.444
Kikuyu		-1.121	-1.433
		(0.157)	(0.087)
		0.326	0.239
Kissi		0.500	-0.223
		(0.402)	(0.663)
		1.649	0.800
Luhya		0.090	-0.044
		(0.825)	(0.916)
		1.095	0.957
Luo		-0.089	0.312
		(0.924)	(0.744)
		0.915	1.367
Ethnic-Level Conformist			
Kalenjin			0.047
			(0.357)
			1.048
Kamba			0.052
			(0.283)
			1.054
Kikuyu			0.126*
			(0.017)
			1.135
Kissi			0.074
			(0.052)
			1.077
Luhya			0.038
			(0.484)
			1.039
Luo			-0.041
			(0.532)
			0.960

* (**) [***] if $p < 0.05$ (0.01) [0.001]

Diagnostics				
RMSE	0.294	0.294	0.332	
J-Statistic			19.046	(0.004)

All specifications include the following individual, household, regional and cluster controls
Individual: age, age^2, dummies for primary, secondary, and higher education.

Individual: dummies for listen to radio, know a place to test for aids
Household: dummies for tv and radio ownership, mortality

Household: dummies for: flooring (tile, wood); roof (tiled)
Cluster Controls: media access, electricity, water and size of cluster

Timing: Age at first marriage, knowledge of contraception

	I	II	III
Dependent variable: CEB			
Ethnicity			
Kalenjin	0.059	0.039	0.196**
	(0.292)	(0.490)	(0.002)
	1.061	1.040	1.216
Kamba	-0.065	-0.082	0.076
	(0.340)	(0.238)	(0.277)
	0.937	0.922	1.079
Kisii	0.017	0.004	0.143
	(0.865)	(0.971)	(0.080)
	1.017	1.004	1.154
Luhya	0.013	0.004	0.239**
	(0.838)	(0.954)	(0.001)
	1.013	1.004	1.270
Luo	-0.053	-0.070	0.280**
	(0.601)	(0.496)	(0.009)
	0.949	0.932	1.323
Religion			
Catholic	-0.005	-0.006	-0.011
	(0.810)	(0.790)	(0.508)
	0.995	0.994	0.989
Muslim	0.011	0.021	0.127
	(0.922)	(0.855)	(0.237)
	1.011	1.021	1.136
No Religion	0.090	0.092	0.078
	(0.082)	(0.082)	(0.221)
	1.095	1.096	1.081
Other Rel	-0.034	-0.034	-0.102
	(0.385)	(0.373)	(0.332)
	0.967	0.967	0.903
Ethnic-Level Education			
Kalenjin		0.212	0.334
		(0.686)	(0.696)
		1.236	1.396
Kamba		0.378	0.066
		(0.698)	(0.962)
		1.459	1.069
Kikuyu		-1.000	-1.236
		(0.227)	(0.224)
		0.368	0.291
Kissi		-0.001	-0.538
		(0.998)	(0.235)
		0.999	0.584
Luhya		-0.172	-0.952*
		(0.696)	(0.037)
		0.842	0.386
Luo		0.437	-0.869
		(0.674)	(0.574)
		1.548	0.420
Ethnic-Level Conformist			
Kalenjin			0.194**
			(0.003)
			1.214
Kamba			0.224**
			(0.003)
			1.251
Kikuyu			0.240***
			(0.000)
			1.271
Kissi			0.162***
			(0.000)
			1.176
Luhya			0.181**
			(0.009)
			1.198
Luo			0.127*
			(0.043)
			1.136

* (**) [***] if $p < 0.05$ (0.01) [0.001]

Diagnostics			
RMSE	0.296	0.297	0.478
J-Statistic			7.493 (0.278)

All specifications include the following individual, household, regional and cluster controls
Individual: age, age^2, dummies for primary, secondary, and higher education.

Individual: dummies for listen to radio, know a place to test for aids
Household: dummies for tv and radio ownership, mortality

Household: dummies for: flooring (tile, wood); roof (tiled)
Cluster Controls: media access, electricity, water and size of cluster

Timing: Age at first marriage, knowledge of contraception

8 Appendix

A I: The Generalised Method of Moments (GMM) estimator

In this paper we adopt the Generalised Method of Moments (GMM) estimator. A GMM estimator is particularly suited to this particular context, since by using first-order conditions only it allows for unobserved heterogeneity without a full set of parametric assumptions (i.e. the Poisson variance assumption), and accommodates endogenous regressors. Whereas the fully parametric mixture extensions of the Poisson regression model, such as the negative binomial and zero-inflated Poisson models, impose statistical independence between the unobserved heterogeneity and $\mathbf{x}_i$, a GMM framework readily accommodates dependence.

The GMM estimator minimises the (*generalized*) length of the empirical vector function

$$\hat{\beta}_{GMM} \equiv \arg\min_{\beta \in \boldsymbol{\Theta}} \psi_N(\mathbf{w}, \beta)' \tilde{Q}_N^{-1} \psi_N(\mathbf{w}, \beta), \tag{4}$$

where $\mathbf{w}$ collects all observables including instrumental variables $\mathbf{z}$. The nature of the conditional moment restrictions, $\psi_N(.)$ will depend upon whether we introduce an additive or multiplicative error structure. As noted by Windmeijer and Santos-Silva (1996), although multiplicative and additive models are observationally equivalent when only the first order conditional mean is specified, there are differences in the presence of endogeneity and the subsequent choice of instruments in the two specifications. Given $E(c_i|\mathbf{x}_i) = \exp(\mathbf{x}_i'\beta)$, the additive model implies $c_i = \exp(\mathbf{x}_i'\beta) + \omega_i$, with moment conditions $E(\mathbf{z}_i'(c_i - \exp(\mathbf{x}_i'\beta))$. The multiplicative model[20] posits $c_i = \exp(\mathbf{x}_i'\beta)\omega_i$, with moment conditions $E(\mathbf{z}_i'(c_i \cdot \exp(-\mathbf{x}_i'\beta) - 1))$. The multiplicative model handles regressors $\mathbf{x}_i$ and error terms ω_i in a symmetric fashion. In addition given that for c not equal to zero, $\ln c_i = \mathbf{x}_i'\beta + \omega_i$ is congruent with the Generalised Linear Model (GLM).

Given a multiplicative error model $\psi_N(\mathbf{w}, \beta) = \mathbf{Z}'(c - \kappa(\beta)/\kappa(\beta))$, $\kappa(\beta) = \exp(\mathbf{X}'\beta)$. We can the write (4) as

$$\hat{\beta}_{GMM} \equiv \arg\min_{\beta \in \boldsymbol{\Theta}} [\frac{1}{n}\mathbf{Z}'(\mathbf{c} - \kappa(\beta)/\kappa(\beta))]' \tilde{Q}_N^{-1} [\frac{1}{n}\mathbf{Z}'(c - \kappa(\beta)/\kappa(\beta))],$$

where $\mathbf{Z} = \{\mathbf{z}_i\}$ denotes an instrument matrix.[21] $\tilde{Q}_N = \mathbf{Z}'\boldsymbol{\Omega}\mathbf{Z}$ denotes the optimal weight matrix where $\boldsymbol{\Omega} = (c - \kappa(\hat{\beta})/\kappa(\hat{\beta}))^2$. $\kappa(\hat{\beta}) = \exp(\mathbf{X}'\hat{\beta})$ with $\hat{\beta}$ a first step GMM estimator based upon any weighting matrix.

A II: Conformist Effects

In the conditional mean specification (3) conformist effects are represented by $\mathrm{cf}_{i,g} = E_i(c_g) - c_i$, where c_i denotes the number of children (CEB) ever born to the ith woman, and $E_i(C_g)$ is the expectation of mean CEB held by the ith woman resident in cluster g. This

---

[20]We note that given the use of a multiplicative model then any additional underlying regression error would be absorbed by the unobserved heterogeneity term ω_i. See Mullahy (1997) for an extended discussion of this point.

[21]The GLM estimator β_{GLM} accounts for both the endogeneity of one or more of the elements of $\mathbf{x}$ and the intrinsic heteroscedasticity of $\kappa_i(\beta) = \exp(\mathbf{x}_i'\beta)$.

formulation contains two components: $E_i(c_g)$ represents the expectation of average fertility of the ith woman in group g. Assuming that each woman possesses rational expectations on $E_i(c_g)$, then individual expectations of average fertility behaviour in the group equals the average behaviour such that $E_i(c_g) = N_g^{-1} \sum_{i \in g} c_i$, where N_g denotes the number of individuals in the group. We argue that the rational expectations is valid given that cluster sizes are small. For example, in Blume, Brock, Durlauf, and Ioannides (2010), it is argued that identification of social interaction effects may depend on whether individuals are reacting to the expected or actual behaviour of others. If the sample sizes are small, then assumptions such as complete information and the individual's payoff depending on the actual choices of others in the group, rather than on expectations about the average choice of others, is justified by the study of interactions in relatively small groups. For example, individuals may have incentives to conform to the behaviour of a small number of appropriately defined neighbours whose actions they can 'observe'.

Below we represent these arguments more formally. Suppose U_i denotes the utility function of household i, c_i represents the number of children, $\mathbf{l}_i$ 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. We extend the standard atomistic utility model of fertility to incorporate a theory of social interactions, so we can write

$$V_i = (U(c_i, \mathbf{l}_i, \mathbf{x}_i); S^j(c_i, \mathbf{c}_{-i}^j)) \tag{5}$$

where V_i, total individual utility, is comprised of a private utility term $U(c_i, \mathbf{l}_i, \mathbf{x}_i)$, and two social utility terms: $S^j(c_i, \mathbf{c}_{-i}^j)$ represents a social utility term for location or ethnicity j, given by some function of the number of children choices made by all other residents of j; $\mathbf{c}_{-i}^j = (c_1^j, c_2^j, ..., c_{i-1}^j, c_{i+1}^j, ..., n_{M_j}^j)$, where M_j denotes the number of individuals in location j. The extension allows us to capture that component of individual utility attributable to a fertility choice c_i, that is dependent upon the fertility choices of others. The 'others' in question may be those, for example, resident in the individual's household, or others resident in the local community for example, in the individual's street or neighbourhood, religion or members of the same ethnic group or tribe. Rewriting (5) as additive in individual and social utility terms, and without loss of generality we have

$$V_i = U(c_i, \mathbf{l}_i, \mathbf{x}_i) + S^j(n_i, \mathbf{c}_{-i}^j) + \varepsilon_i. \tag{6}$$

The addition of the social utility function $S^j(.)$ represents a model in which we allow individual behaviour to be partially determined by the collective decisions of a group. Differentiating the Beckerian first-order conditions with respect to social interactions yields the second-order conditions

$$\kappa_{il}^j = \partial^2 S^j(c_i, c_l)/\partial c_i \partial c_l > 0,$$

where κ_{il}^j represents a measure of the disutility accruing to i from deviating from the behaviour of l for $i, l \in j$.

The form of social utility at the level of location j may be written as

$$S^j(c_i, \bar{c}^j_{-i}) = \frac{-\theta_2}{2}(c_i - \bar{c}^j_{-i})^2, \tag{7}$$

where (7) represents social utility as a measure of conformism with $\partial S(.)/\partial c_i = \theta_2(\bar{c}^j_{-i} - c_i)$. Under (7) there is an incentive to conform to mean fertility behaviour within the group with deviations from the mean penalised. If behaving like others confers additional status on a woman, she may desire to conform.[22]

A III: The Identification Problem

The nature of the identification problem in models of social interaction is complex. Below we consider a number of dimensions to this problem, namely: the presence of timing effects, Manski's reflection problem, the endogeneity of mean effects, and identification of interaction effects. Such timing effects, for example younger cohorts marrying at later ages, are important in that they determine, in part, the distance between the observed number of children for any given woman and the mean number of children in the appropriate reference group. This difference is central to our thesis, namely that changes in this distance will impact upon CEB due to an interaction effect. Below we address each of these issues, and in doing so outline our identification strategy.

In considering the relevance of the reflection problem for our particular application, we first outline the identification problem in the linear-in-means social interaction model as originally considered by Manski (1993), and more recently by Blume, Brock, Durlauf, and Ioannides (2010) and Ioannides (2008). Since the question of identification is well understood in the linear framework, this will provide a useful point of departure. Using what it now common notation for this problem, a linear model of fertility behaviour may be written as

$$c_i = \alpha + \mathbf{q}'_i\eta + \mathbf{y}'_g\lambda + \kappa m^e_{ig} + \varepsilon_i, \tag{8}$$

where $\mathbf{q}_i$ denotes a $K \times 1$ vector of individual characteristics, $\mathbf{y}_g$ is a $L \times 1$ vector of exogenous (contextual) characteristics and m^e_{ig} denotes the expected average fertility of others in the group held by the ith woman. η, λ, and κ denote unknown parameters. Under a rational expectations assumption we may write $m^e_{ig} = m_g = \alpha + \bar{\mathbf{q}}'\eta + \mathbf{y}'_g\lambda$. If we then impose the restriction that the dimension of $\bar{\mathbf{q}}$ is the same as $\mathbf{y}_g$, then m_g may be written as

$$m_g = \frac{\alpha + (\eta + \lambda)\mathbf{y}_g}{1 - \kappa}. \tag{9}$$

In this context Manski's reflection problem is immediately seen. Namely, any observed correlation between an outcome c_i and m_g solely reflects the combined effect of mean-group characteristics, $\mathbf{y}_g$, rather than endogenous interaction within a group. The genesis of the reflection problem derives from the fact that in (8) we specify a model for children ever born which is linear in an endogenous effect, m^e_{ig}, a vector of exogenous effects, $\mathbf{y}_g$,

---

[22]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.

and a vector of individual characteristics, $\mathbf{q}_i$. If the number of exogenous and individual effects are equivalent it then follows that the reduced form representation of the conditional mean will be a function of $\mathbf{y}_g$.[23]

There are a number of ways to circumvent the reflection problem. Within a linear framework the simplest method is to impose a full set of zero restrictions on λ, thereby adopting the assumption that any interactions are mediated solely by endogenous effects. We note that such an approach is potentially problematic for making inference on multiplier effects, since such restrictions may result in biased estimates of pure endogenous effects (see Ioannides (2008)). This is important since policy proscriptions which follow from endogenous versus exogenous effects are likely to be very different. However, the upside of such an approach is that these restrictions automatically generate a possible set of instruments ($\mathbf{q}$) for m_g.[24] Second, and as discussed earlier, identification is facilitated by locating one or more individual level variables whose group level average does not enter the conditional mean.

In this study the reflection problem is mitigated for a number of reasons. First, although 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. Numerous studies in developed and developing countries have argued that education is the single most important predictor of fertility decline. This is because education increases the opportunity cost of women's time, and encourages parents to have higher quality children, which then reduces women's fertility. Moreover, better educated women are likely to exert an influence on less educated women if they provide effective role models for them. It is for this reason that we focus primarily on this ex ogenous effect. Second, given that the conditional mean specification represents a nonlinear transformation of the endogenous effect, then even if the restriction $\overline{\mathbf{q}} = \mathbf{y}_g$ held, identification would be achieved.

Endogeneity Bias

Given the form of endogenous conformist effects, namely $\mathrm{cf}_{i,g} = E_i(c_g - c_i)$, we face a variant of the the mean endogenous effects problem. In tackling this problem we adopt an instrumental variable approach. However, the required condition that an instrument may only impact c_i indirectly through its association with $\mathrm{cf}_{i,g}$ may not hold given that $\mathrm{cf}_{i,g}$ is a linear function which includes c_i.

We address this problem by an appropriate specification of the conditional mean. Given that the mean of a count process has non-negative support, we specify conditional mean of CEB as an exponential function of a linear index in the regressors. As a consequence, the otherwise linear dependence between the original error term, ω_i, and the instrument is removed given the imposition of a restriction on the support of the conditional mean using the exponential function[25]. It is important to emphasise that the nature of our

---

[23]See Manski (1993) and Blume, Brock, Durlauf, and Ioannides (2010) for further details.

[24]See, for example, Case and Katz (1991) and Gaviria and Raphael (2001). See Kawaguchi (2004) for an approach which utilises subjective perceptions of peers' average behaviour, instead of average behaviour.

[25]Ioannides (2008) in an extensive overview of the empirics of social interactions also notes that nonlinearities may be useful in providing identification in models of social interactions.

identification strategy is *not* dependent on a highly specific set of parametric assumptions but rather a minimal set of assumptions to ensure the non-negativity of the conditional mean. Blume, Brock, Durlauf, and Ioannides (2010) make a similar argument in outlining conditions for identification in other nonlinear social interaction models.[26].

Given that the number of instruments excluded exceeds the number of endogenous variables, our estimating equation is overidentified, and we may evaluate the validity of our instruments by testing whether the excluded instruments are independent of the error process using the J-test (see Hansen (1982)). This statistic is the value of CEB objective function, evaluated at the efficient CEB estimator. Under the joint null hypothesis of correct model specification and the validity of the orthogonal conditions, the J-statistic is given by

$$J(\widehat{\beta}_{\text{CEB}}) = n\mathbf{Z}'\widehat{\mathbf{u}}(\mathbf{Z}'\widehat{\Omega}_c\mathbf{Z})^{-1}\mathbf{Z}'\widehat{\mathbf{u}} \sim \chi^2_{l-k},$$

where $\widehat{\mathbf{u}}$ denotes the residuals and l-k denotes the number of overidentifying restrictions. $\widehat{\Omega}_C$ represents a form of the error covariance matrix which is robust to arbitrary intra-cluster correlation.[27] A rejection of the null may then result from the instruments not satisfying the orthogonality conditions given they are not exogenous or because they are incorrectly excluded.

---

[26]See Brock and Durlauf (2003) for a discussion of identification in the multinomial choice model with social interactions.

[27]Note that cluster robust standard errors are asymptotic in the number of clusters (M). Kezdi (2004) shows for a given type of model that M=50 will generate reasonable inference (see White (1984)). Baum et al (2003) provide a useful discussion of the construction of the $\widehat{\Omega}_C$ matrix.