KU LEUVEN

FACULTY OF BIOSCIENCE ENGINEERING

Fertility preferences and the trade-off between quantity and quality of child-raising: insights from southern Ethiopia.

Fertiliteitpreferenties en de trade-off tussen kwantiteit en kwaliteit in familie-economie: een studie in zuidelijk Ethiopië.

Promotors:
Prof. Miet Maertens
Department of Earth and Environmental Sciences
Division of Bioeconomics

Prof. Peter Goos
Department of Biosystems
Division of Mechatronics, Biostatistics and Sensors

Dissertation presented in
Fulfillment of the requirements
for the degree of Master of Bioscience Engineering:
Agro- and Ecosystems Engineering

Eva Boonaert

June 2020

Acknowledgments

In the first place, I would like to express my gratitude to my promotor Prof. Dr. Ir. Miet Maertens who gave me the opportunity to delve into this really interesting topic and who taught me with great enthusiasm about this research domain. Next, I would like to thank my previous copromotor Prof. dr. Ir. Goedele Vanden Broeck and current copromotor Prof. dr. Peter Goos for the regular and constructive feedback sessions. A special thanks to Prof. dr. Peter Goos showing me the joy in experimental design. Moreover, I would like to thank Kaat Van Hoyweghen. You were an excellent mentor and were always ready to offer help, give good advice and for reading all my texts.

Further, I would like to thank the team members of the VLIR-UOS project who gave me the opportunity to be part of this interesting project. My biggest thank you goes to Ashenafi Duguma Feyisa, for his tremendous help during the fieldwork with the practical arrangements and outside the fieldwork as well for the nice time we had and to get to know me the real local culture. It was a pleasure to work with you! Next, I would like to thank Aladar for his help with the tablets, Mule for his good drive skills to ride even the most unpaved trails and my enumerators (Tsige, Tamrat, Hakime, Amanuel, Markos and Merawi) for their commitment to conduct the surveys which was not always that easy due to the different languages and remote villages; and the fun time we had during the fieldwork. Furthermore, a big thank you to Lisa, Margot, Simon, Lana, Maxime, Monica, Nansol and Sara for the nice time during our stay in Arba Minch and sharing the struggles searching for internet to work together on our thesis.

Most of all, I would like to praise my parents for their enormous support during my educational journey. There are no words to describe how much I appreciate what you have done for me. I am forever grateful! Special thanks also for rereading my thesis.

Lastly, I would like to thank my brother for his support, my boyfriend for being always there for me, my roommates for all the cosy evenings, kotmeals, parties and kotweekends, and my friends from the youth movement, high school and bio-science engineering, in particular Léonie, Eva, Alessia, Emma, Eline, Hanne, Inès, etc., friends I could build on.

I hope you enjoy reading it!

ሁላችሁንም አመሰግናለሁ! (Thank you all)

Abstract

Sub-Saharan Africa is facing high population growth related to the high total fertility rate. This thesis performs a case study questioning the variables related with higher fertility preferences, the existence of a quantity-quality trade-off (where quality refers to education in this study) and the possible sex preference in both quantity and quality of child-raising. This study is innovative since a discrete choice experiment with a non-linear mixture-amount model was implemented, focussing on ex ante fertility preferences (i.e. before completing their reproductive lifetime) in six rural districts in southern Ethiopia. This was complemented with quantitative follow-up questions and qualitative focus group discussions.

Our results demonstrate three elements. First, the choice experiment revealed that the latent preference for the ideal family size is on average 5.98 children for men and 5.62 children for women. This is more than one child higher than the self-expressed preferred family size during the follow-up questions. This gap is explained by the higher family size preference per se and not by an unmet need for contraception. Moreover, there is a large heterogeneity in the preferred number of children. The main determinants related with a higher family size preference are associated with the cultural diffusion theory, and the economic and investment theory, namely *self-expressed family size preference, being married, having children, being male, place of residence* (i.e. living in the Gamo Gofa Zone or Konso district) and *belonging to the Konso ethnic group*. Recommendations for family planning policies are to focus, next to contraception, on decreasing the family size norm of these target groups and to increase birth registration. Second, our empirical analysis confirms the existence of a quantity-quality trade-off for most educational levels. Surprisingly, in most cases, this only holds until the level at which respondents reach their preferred family size. This reflects that respondents are willing to give up schooling in favour of having more children until reaching their preferred family size level. After reaching their preferred family size, respondents are only willing to have more children if these children can attain a higher educational level. The existence of a quantity-quality trade-off implies that increasing the availability, accessibility and affordability of education can stimulate declines in quantity. Third, in general, men have a sex preference for sons concerning the preferred number of children and investment in education. To reduce the gender inequality in education, mainly men need to be targeted.

KEYWORDS: Fertility, Education, Gender, Sex preference, Quantity-quality trade-off, Mixture-amount model, Choice experiment, Ethiopia

Samenvatting

Sub-Sahara Afrika wordt geconfronteerd met een hoge bevolkingsgroei als gevolg van het hoge totale vruchtbaarheidscijfer. Via een casestudy onderzocht deze thesis de variabelen gerelateerd aan hoge vruchtbaarheidspreferenties, de afweging tussen kwantiteit en kwaliteit op het vlak van onderwijs en het mogelijke verschil in gendervoorkeur. Om deze vragen te beantwoorden werd een keuze-experiment met een niet-lineair mixture-amount model uitgevoerd om ex ante vruchtbaarheidspreferenties te onderzoeken (d.w.z. voor het bereiken van de menopauze) in zes rurale districten ten zuiden van Ethiopië. Dit werd aangevuld met kwantitatieve vragen en focusgroepsdiscussies.

De resultaten belichten drie elementen. Ten eerste toont het keuze-experiment aan dat de latente ideale gezinsgrootte gemiddeld 5.98 kinderen bedraagt voor mannen en 5.62 kinderen voor vrouwen. Dit is meer dan één kind hoger dan de zelfverklaarde ideale gezinsgrootte. Deze kloof wordt verklaard door de voorkeur voor een hogere gezinsgrootte an sich en niet door een onvervulde behoefte aan anticonceptie. Bovendien is er een grote heterogeniteit in de vruchtbaarheidsvoorkeuren. De belangrijkste determinanten inzake een hoger geprefereerde gezinsgrootte zijn gerelateerd aan de culturele diffusietheorie, en de economische en investeringstheorie, namelijk *ideale gezinsgrootte*, *getrouwd zijn*, *kinderen hebben*, *man zijn*, *woonplaats* (i.e. wonen in de Gamo Gofa Zone of Konso-district) en *behoren tot de Konso-etnische groep*. Aanbevelingen voor het huidige beleid zijn om, naast anticonceptie, te focussen op het verkleinen van de gewenste gezinsgrootte van deze doelgroepen en het verhogen van de geboorteregistratie. Ten tweede bevestigt onze empirische analyse het bestaan van een afweging in kwantiteit en kwaliteit voor de meeste onderwijsniveaus. Opvallend is dat dit in het algemeen enkel geldt tot het punt waarbij de ideale gezinsgrootte is bereikt, wat aangeeft dat de respondenten bereid zijn om onderwijs op te geven om meer kinderen te krijgen. Na het bereiken van de ideale gezinsgrootte zijn respondenten enkel bereid om meer kinderen te krijgen als deze een hoger opleidingsniveau kunnen behalen. Het bestaan van een afweging in kwantiteit en kwaliteit impliceert dat een grotere beschikbaarheid, toegankelijkheid en betaalbaarheid van onderwijs dalingen in kwantiteit kunnen stimuleren. Ten derde hebben mannen over het algemeen een gendervooroordeel inzake aantal kinderen en onderwijs. Om de genderongelijkheid in het onderwijs te verkleinen, moet het beleid vooral op mannen focussen.

TREFWOORDEN: Vruchtbaarheid, Onderwijs, Gender, Gendervoorkeur, Kwantiteit-kwaliteitsafweging, Mixture-amount model, Keuze-experiment, Ethiopië

List of abbreviations

ABE	Alternative Basic Education
AICc	Akaike information criterion corrected for small sample sizes
ANA	Attribute Non-Attendance
BIC	Bayesian information criterion
CLM	Conditional logit model
DCE	Discrete choice experiment
EAAlogit	Endogenous attribute non-attendance logit model
FSP	Family size preference
LC	Latent class
LCM	Latent class model
LL	Log-likelihood
MNLM	Multinomial logit model
MRS	Marginal rate of substitution
MXLM	Mixed logit model
NGO	Non-governmental organisation
SSA	Sub-Saharan Africa
SNNPR	Southern Nations, Nationalities and People’s Region
TFR	Total fertility rate
TVET	Technical and vocational education and training

List of figures

List of tables

Table of Contents

1. Introduction

The exceptionality of Sub-Saharan Africa's population dynamics has been a key issue in much of the fertility literature. While the world population is predicted to approach 9.8 billion in 2050 and reach 11.2 billion by 2100, more than half of the growth towards 2050 is expected to come from Africa (Figure 1-1) (The World Bank, 2017). After 2050, Africa is predicted to be the only continent which still increases in population because of the young age structure, mainly in Sub-Saharan Africa (SSA). This region is characterized by rapidly declining mortality rates whereas their fertility rates are the highest in the world and are declining at a slower rate compared to other countries. Analysis of the population dynamics throughout the world is necessary to make predictions for the future. One of the most influential theories which tries to do this is the demographic transition model (Thompson, 1929). This model shows that population dynamics are influenced by two factors (making abstraction of any migration), namely mortality and fertility. Mortality rates drop due to scientific and technological progress, which is in general a sign of development that one does not want to slowdown for the sake of limiting population growth (Perman et al., 2011). Consequently, the major possible control on the current population growth is through fertility control.

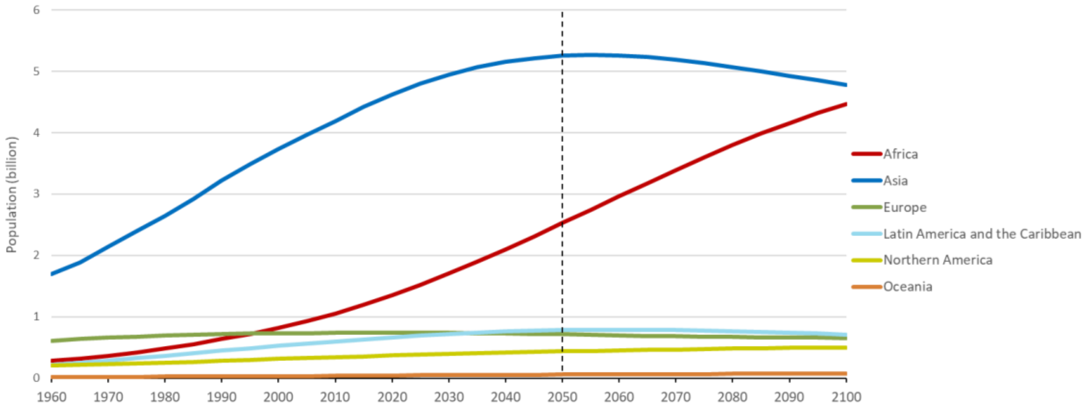


Figure 1-1: Probabilistic population projections per continent, 1960-2100 (The World Bank, 2020; United Nations, 2019b)

Fertility control can be achieved in different ways e.g. via increasing access to birth control, education, female empowerment, economic development, influencing social norms, stimulating further declines in mortality rates, etc. (Atake & Gnakou Ali, 2019; Shenk et al., 2013). Insights in which of these factors are related with higher fertility preferences are needed to improve current family planning policies and programs in order to stimulate the slowdown in population growth.

Persisting knowledge gaps exist about fertility preferences in SSA. Literature concerning fertility preferences reveals three existing voids. First, there is consensus about seven direct determinants influencing fertility whereas there is contentious evidence about the indirect factors (Matthijs, 2012). Three models try to explain the indirect variables affecting fertility: a change in attitudes due to socio-cultural influences (i.e. the cultural diffusion theory), insurance behaviour (i.e. risk and mortality theory) and human capital and wealth (i.e. the economic and investment theory) (Shenk et al., 2013; Werding et al., 2014). Nevertheless, no one can claim full evidence. Second, it is often argued that, as populations undergo the demographic transition, incentives to decrease the number of children (i.e. quantity) and rise the investment in human capital (i.e. quality which refers to education in this study) increase (Galor, 2012). A wide range of economic theories have been proposed concerning this quantity-quality trade-off whereas empirical evidence in SSA in scarce and inconclusive (Clarke, 2018). However, it is important to understand to what extent there is a trade-off between quantity of children and their education; and whether this is rooted in preferences. The reason is that there is a correlation between education of parents and their fertility (KC & Lutz, 2014). If it would turn out that high fertility rates and low schooling rates go hand in hand, future generations will be less educated and have more children; and therefore reaching lower population growth will be difficult. Third, both the preference for investment in the number and human capital of children can be gender biased (de la Croix & Perrin, 2016; Galor, 2012). It is important to investigate this, since, if girls are less educated than boys, female empowerment in future generations will remain low which will again lead to more children and complicate the achievement of lower population growth (Atake & Gnakou Ali, 2019). Recent empirical evidence concerning the gender gap in education in SSA reveals that there is some evidence for a preference for boys in education (Kazeem et al., 2010; Kuépié et al., 2015; Vimefall et al., 2017). However, empirical research concerning the gender gap in preference for sons or daughters in SSA is limited and recent studies show mixed evidence (Chao et al., 2019; Fuse, 2010; Rossi & Rouanet, 2015).

Therefore, this research focuses on the following three research questions:

1. Which factors explain the heterogeneity in fertility preferences?
2. Do rural households make a trade-off in the quantity and quality of child raising, especially education?
3. Are fertility preferences and the quality quantity trade-off gendered?

This thesis tries to fill the gap in the existing literature by conducting a case study in Ethiopia. A discrete choice experiment complemented with follow-up questions was implemented in six rural districts of the Segen People's and Gamo Gofa Zones in southern Ethiopia. 426 respondents, between the age of 18 and 25, were interviewed. Ex ante fertility preferences (i.e. before completing their reproductive lifetime) were studied because of two reasons. First, a focus on ex ante preferences was chosen to minimize bias due to rationalization because of having already a certain number of children (Bongaarts, 2011). The second reason was that fertility preferences are shown to vary significantly over individual's lifetime due to high existential uncertainty in SSA (Trinitapoli & Yeatman, 2018). In addition, focus group discussions were held before the choice experiment to inform the design of the experiment. Afterwards, the data was analysed using the econometric mixed logit and latent class models accounting for possible attribute non-attendance and scale heterogeneity.

Why Ethiopia? This is a relevant case for two reasons. First, Ethiopia is the 12th most populated country in the world in combination with a low educational completion rate (completion rate of 54% in primary, 29% in secondary and 8% in tertiary education) (The World Bank, 2020; United Nations, 2019b). Second, family planning policies and interventions focusing on improving the access to contraception during the last 30 years resulted in a steep decrease in the total fertility rate (TFR) from 7.2 to 4.1 births per woman (The World Bank, 2020; Towriss & Timaeus, 2018). Nevertheless, the aimed reduction in fertility rates is not achieved yet.

This study is innovative in three ways. First, the analysis of fertility preferences was done via a discrete choice experiment (DCE) which is quite uncommon to study fertility preferences. This can reveal the stated family size preference (FSP) in another way than the self-expressed ideal or desired FSP (Bongaarts, 2011; Channon & Harper, 2019). By revealing respondent's latent preferences, the hypothetical and social desirability bias can be omitted. Second, a mixture-amount model was used which is a useful model to study the quantity-quality trade-off due to the constraint that the total number of children attaining a certain educational level need to sum up to the total number of children. However, this model is still quite uncommon in choice experiments, despite the work of some authors (Goos & Hamidouche, 2019; Khademi & Timmermans, 2012; Pradhan et al., 2017; Raghavarao & Wiley, 2009; Ruseckaite et al., 2017). Third, our contribution is that we assumed a non-linear utility function, as we assume that, when the quantity changes, this will change respondents' preferences for quality. To our knowledge, only one study has done this before by Mogstad and Wiswall (2009).

In the following chapters, first a literature overview concerning fertility preferences, the quantity-quality trade-off, gender differences and a focus on the situation in Ethiopia is given. This is followed by a description of the methodology about the data collection. Subsequently, the results are shown and discussed. Afterwards, a conclusion and policy recommendations are represented.

2. Literature review

2.1. Measurement of fertility

In this research, fertility is defined as "the number of children born alive that women give birth to (excl. stillbirth, foetal death and abortion)" (Matthijs, 2012). Fertility can be measured in distinct ways. At the societal level, the most realistic measure is the total fertility rate (TFR). This is a hypothetical measure for the average number of children that would be born per woman according to the same fertility behaviour across the population until the end of her reproductive life (assumed from 15 to 49 years) as determined in a given year (Hinde, 1998) (Eq. 1):

Eq. 1 $$\mathrm{TFR_t} = \sum_{x=15}^{49} \frac{number\ of\ births\ in\ year\ t\ to\ women\ aged\ x\ to\ x+1}{\text{mid} - \text{year population of women aged x to x} + 1}$$

This is a more direct measure of fertility than the crude birth rate[1], the general fertility rate[2] or the bruto[3]/netto[4] reproduction rate since it is independent of the age structure of the population. At the individual level, the ideal or desired family size is the least problematic measure as explained by Bongaarts (2011). The ideal or desired FSP is asked as: "If you could go back in time to the moment when you did not have any children yet, how many children would you choose?". Other possible measures to assess childbearing preferences are the wanted fertility[5], the wanted status of recent births[6] and desire for more children[7] (Bongaarts, 2011; Channon & Harper, 2019).

2.2. The demographic transition model

Fertility decline in human history is a complex evolutionary puzzle. Several models try to explain this. The most popular is the demographic transition model as proposed by Thompson (1929) and further elaborated by Notestein (1945), Landry (1987) and Pritchett (1994). This model describes the transition from high birth and death rates (the first or pre-transitional stage) to lower birth and death rates (the fourth or post-transitional stage) and the corresponding population increase (Figure 2-1). This model shows that population growth is affected by two main factors (making abstract of any migration), namely mortality (which is inversely proportional with the life expectancy) and fertility (Thompson, 2003).

---

[1] Measured as the total number of births in year t divided by the total mid-year population (Hinde, 1998).
[2] Measured as the total number of births in year t divided by the total mid-year population of women of childbearing age (Hinde, 1998).
[3] Measured in the same way as the TFR but only counting the number of girls which are born (Rowland, 2010).
[4] Measured in the same way as the TFR but only counting the number of girls which are born and taking into account the age-specific mortality rates as well (next to the age-specific fertility rates) (Rowland, 2010).
[5] Measured in the same way as the TFR but only the number of observed births before the desired family size is reached (i.e. wanted births) are used in the numerator (Bongaarts, 2011).
[6] Measured by asking women whether recent births are wanted (Bongaarts, 2011).
[7] Measured by asking women whether they want more children (Bongaarts, 2011).

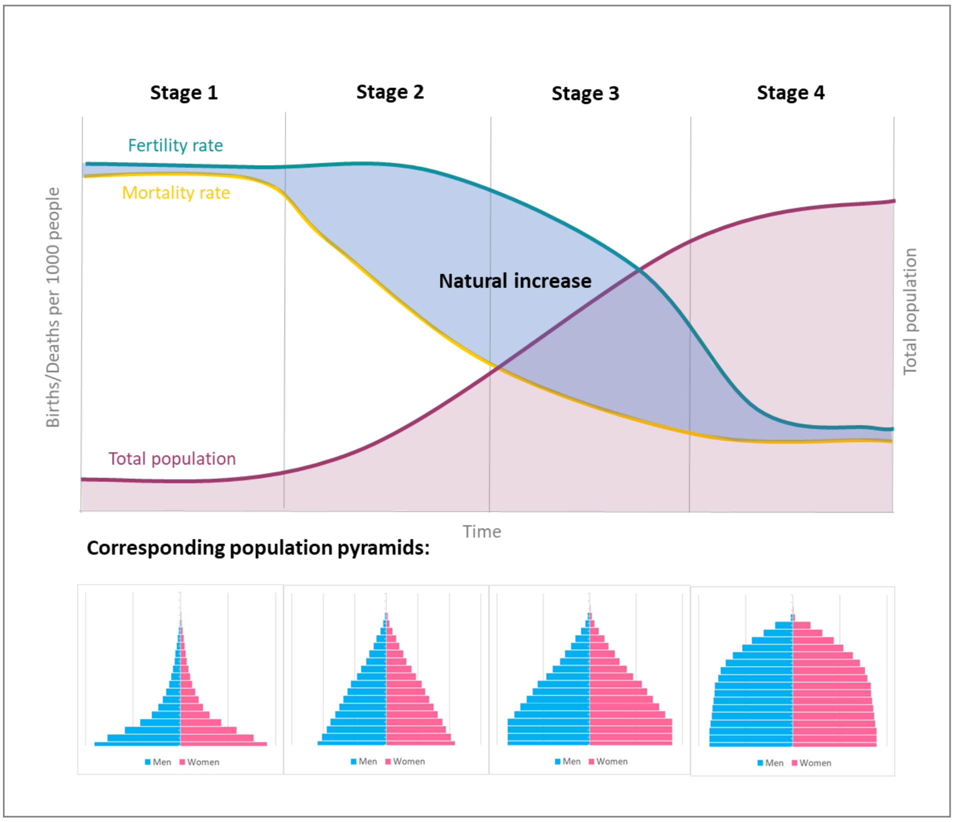


Figure 2-1: Visualization of the demographic transition model with the three determinants: the mortality rate, the fertility rate and the total population (adapted from Roser et al. (2019))

The demographic transition model is based on the observed historical population dynamics of **western countries** (Perman et al., 2011)**.** These are now assumed to be in stage four. Analysis of their population growth can be used to make predictions for the future for other countries passing through the demographic transition. The slow reduction in mortality rates in stage two (i.e. the early transitional stage) coincided with the start of the industrial revolution (Galor, 2012). This period is characterized by a process of modernization, urbanization and industrialization. In the third stage (i.e. the late transitional stage), the birth rates also started to decline slowly. Analysis of the demographic transition of western countries reveals two elements. First, the decline in mortality rates, which always occurs before the decline in fertility rates, is mainly a result of scientific and technical developments (mainly improved food provision, health, hygiene, education, government institutions and diminishing influence of religion). Second, the decline in fertility rates is shown to be related to a wider range of variables. Therefore, fertility decline is more difficult to predict (Gerland et al., 2014).

Nevertheless, for many **developing countries,** the industrial revolution started with a delay. These countries are now assumed to reside in stage two or three (Perman et al., 2011). They have reached the second stage as a consequence of knowledge and technology transfer resulting in a much faster decline in mortality rates than had occurred in the developed countries. However, the fall in birth rates is lagging behind. This is mainly the case for SSA (The World Bank, 2020).

While the crude death rate[8] in SSA decreased by 59% in the past 50 years, the total fertility rate (TFR) is the highest in the world with 4.8 births per woman, compared to 2.4 as the world average and it is declining at a slower rate compared to developing countries in Asia and Latin-America (Bongaarts & Casterline, 2013; The World Bank, 2020). The wider gap between the decline in mortality and fertility rates for developing countries compared to developed countries explains the current higher population increase in developing countries (The World Bank, 2017).

Mortality rates drop due to scientific and technological progress, which is in general a sign of development that one does not want to slowdown for the sake of limiting population growth (Perman et al., 2011). Consequently, the major possible control on the current population growth is through fertility control. Therefore, insights in the factors affecting the decline in fertility rates are needed to understand the observed population dynamics in SSA and to improve current family planning policies and programs.

2.3. Factors influencing the decline in fertility rates

Fertility is defined by various determinants. Davis and Blake (1956) were the first to make a distinction between direct and indirect determinants. Based on this work, Bongaarts (1978, 2015, 2017) proposed a framework to quantify the fertility effects (Figure 2-2). This way, the mechanisms through which the factors have an impact on fertility can be analysed. In the next paragraphs, the empirical evidence for the different determinants is analysed focussing on the following four criteria: i) recent (2010-2020), ii) peer-reviewed articles on Web Of Science and Scopus or referred sources from these articles, iii) focussing on rural areas, iv) in Sub-Saharan Africa and Ethiopia.

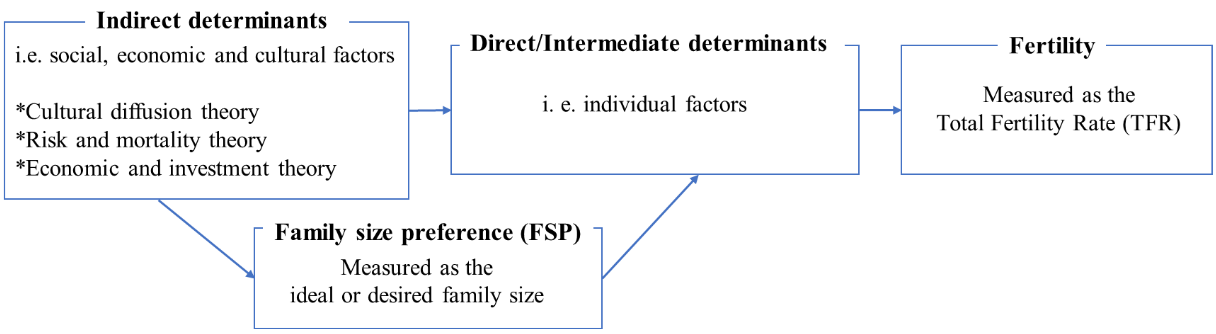


Figure 2-2: Conceptual framework of Bongaarts' fertility model (1978)

[8] Measured as the number of deaths occurring at midyear per 1000 population (The World Bank, 2020).

2.3.1. Direct determinants

There is consensus among demographers about seven main direct factors affecting fertility (Bongaarts, 1978, 2015, 2017; Bongaarts & Casterline, 2013; Bongaarts & Potter, 1983; Matthijs, 2012). Empirical research in SSA indeed reveals that *the proportion of women who are married* (Ariho & Kabagenyi, 2020; Burger et al., 2012; Defo, 2011; Kodzi et al., 2012; Rutaremwa et al., 2015) or more topical *the proportion of women who is sexually active* (Ariho & Kabagenyi, 2020; Defo, 2011) and *the frequency of sexual intercourse* (Chemhaka & Odimegwu, 2019) increase fertility, whereas *spacing* (Timæus & Moultrie, 2020), *contraceptive use* (Ariho & Kabagenyi, 2020; Chemhaka & Odimegwu, 2019; Colleran & Snopkowski, 2018; Rutaremwa et al., 2015), *induced and spontaneous abortion* (Chemhaka & Odimegwu, 2019), *postpartum insusceptibility* (i.e. postpartum abstinence or postpartum amenorrhea) (Chemhaka & Odimegwu, 2019; Rutaremwa et al., 2015) and *sterility* (related to the start of the menarche, menopause and biological sterility for which there is no recent evidence) have a negative effect on fertility. Further, c*ontraception* is found to be the most important factor influencing fertility in most researches.

Focussing on Ethiopia, previous studies reveals the same main factors affecting fertility. *The proportion of women who are married* (Alemayehu et al., 2010; Laelago et al., 2019; Mekonnen & Worku, 2011) and *the proportion of women who is sexually active* (Berlie & Alamerew, 2018) positively influence fertility. Furthermore, there is also evidence that *induced and spontaneous abortion* (Laelago et al., 2019), *postpartum insusceptibility* (Alemayehu et al., 2010; Laelago et al., 2019; Teklu et al., 2013) and *spacing* (Towriss & Timaeus, 2017) negatively influence fertility. The same holds for c*ontraceptive use* (Alemayehu et al., 2010; Alvergne et al., 2013; Aragaw, 2015; Teklu et al., 2013; Wado et al., 2019) for which Ali et al. (2012) report high levels of contraceptive discontinuation. No recent studies are found for the factors *frequency of sexual intercourse* and *sterility*. Moreover, *marriage* is identified as the most important determinant in urban areas, whereas *postpartum insusceptibility* is the major variable in rural regions (Alemayehu et al., 2010).

2.3.2. Indirect determinants

Concerning the indirect determinants, three models try to explain the mechanisms through which the indirect factors affect the FSP and the direct factors: a change in attitudes due to socio-cultural influences (i.e. the cultural diffusion theory), insurance behaviour (i.e. risk and mortality theory) and human capital and wealth (i.e. the economic and investment theory) (Shenk et al., 2013; Werding et al., 2014).

2.3.2.1. The cultural diffusion theory

Concerning socio-cultural influences, the diffusion theory is the most well-known, first proposed by Leroy-Beaulieu (1896). This states that ideas and behaviours about fertility and family planning are diffused through mass media, opinions of leaders, social network, etc. Empirical evidence in SSA shows that *self-esteem* (Ibrahim & Arulogun, 2019), *peer-related subjective norms* (Ibrahim & Arulogun, 2019), *perceived parental expectation of fertility* (Ibrahim & Arulogun, 2019), *polygamy* (Baschieri et al., 2013) and *gender of the respondent* (Ibrahim & Arulogun, 2019) are linked with fertility decline. However, there is mixed evidence about *religion* (Hayford & Agadjanian, 2011; Ibrahim & Arulogun, 2019) and the *age group* (Ibrahim & Arulogun, 2019; Kodzi et al., 2012). Further, *media* and *ethnicity* seem to have no effect (Ibrahim & Arulogun, 2019).

Recent evidence from Ethiopia finds the following determinants: *access to media* (Ferede, 2013), *age at first birth* (Berlie & Alamerew, 2018), *age group* (Alemayehu et al., 2010; Ferede, 2013), *ethnic group* (Teklu et al., 2013), *religion* (Teklu et al., 2013) and *family-related subjective norms* (Teklu et al., 2013).

2.3.2.2. The risk and mortality theory

The risk and mortality model states that reduced risk diminishes the number of births needed to realize the desired family size (Bongaarts, 2011; Coale, 1973; Galor, 2012). Risk can be related to infant or child mortality or high levels of stress. Recent evidence from SSA reveals that a reduction in *infant/child mortality* (Eastwood & Lipton, 2011; Kodzi et al., 2012; Shapiro & Tenikue, 2017) and *climatic variability* (Eissler et al., 2019) decreases fertility. In addition, it is observed that, for all countries, the decline in mortality rates preceeds the decline in fertility rates (Galor, 2012). In Ethiopia, authors report that a *history of child death* and *food insecurity* is related with a high FSP (Mekonnen & Worku, 2011).

2.3.2.3. The economic and investment theory

Concerning the economic and investment theory, fertility decline is a result of a change in costs and benefits of child-raising due to a change in human and financial capital (Easterlin, 1975; Livi Bacci, 2017; Pritchett, 1994). Empirical evidence in SSA finds that: *women's education* (Atake & Gnakou Ali, 2019; Buyinza & Hisali, 2014; Chemhaka & Odimegwu, 2019; Eastwood & Lipton, 2011; Shapiro & Tenikue, 2017), *education of both spouses* (Burger et al., 2012; Colleran & Snopkowski, 2018; Rutaremwa et al., 2015, whereas, Channon and Harper (2019) do not find significant evidence), *women's bargaining power* (Atake & Gnakou Ali, 2019; Moya et al., 2016, although Upadhaya and Karasek (2012) do not find consistent evidence), *women's control over household resources* (Atake & Gnakou Ali, 2019), *women's participation in income-generating activities* (Atake & Gnakou Ali, 2019) and *place of residence (urban vs rural)* (Rutaremwa et al., 2015) have a negative impact on fertility. However, the positive effect of *household income or economic growth on country-level* (Colleran & Snopkowski, 2018; Kodzi et al., 2012; Rutaremwa et al., 2015) as well as the *old-age security* hypothesis (Root & Johnson-Hanks, 2016) is refuted by Galor (2012). Furthermore, it is observed that most of the studies analysing the effect of indirect determinants focus on the link with contraception which affects in turn the TFR.

Empirical studies in Ethiopia reveal that the main factors decreasing the TFR are: *women's education* (Berlie & Alamerew, 2018; Ferede, 2013; Mekonnen & Worku, 2011; Towriss & Timaeus, 2017), *education of both spouses* (Alemayehu et al., 2010; Behrman, 2015), *women's bargaining power* (Wado et al., 2019), *place of residence* (Alemayehu et al., 2010; Ferede, 2013; Wado et al., 2019), *access to family planning information and services* (Ferede, 2013; Wado et al., 2019), *employment* (Alemayehu et al., 2010) and *knowledge about contraception* (Wado et al., 2019). However, no effect is found for *contraceptive prevalence* (Desai & Tarozzi, 2011).

Overall, there is contentious evidence about which factors and mechanisms explain the decreasing course of fertility in the demographic transition model (Gerland et al., 2014). Insights in the factors explaining the heterogeneity in fertility preferences in SSA are key to improve current family planning policies and programs. If for example fertility preferences seem to be more a question of social and cultural norms, policies based on incentives to decrease risks or increase economic or human capital investments have little impact (de la Croix & Perrin, 2016).

2.4. The quantity-quality trade-off

2.4.1. Theoretical evidence

It is often argued that as populations undergo the demographic transition, incentives to decrease the number of children and increase the investment in human capital increase (Galor, 2012). This concept of 'quantity-quality trade-off' originates from economics (namely the cultural-diffusion and economic and investment theories) and evolutionary ecology (Lawson & Borgerhoff Mulder, 2016). Quality is defined in this research as "child-resources which are devoted to human-capital augmenting education" (Fernihough, 2017).

In evolutionary ecology, the quantity-quality trade-off originates from the fact that resources (e.g. energy, time, etc.) are finite, first defined by Lack (1954). Natural selection results in a maximization of this trade-off, which rationales the assumption of individual's utility maximizing behaviour in economics. In combination with the idea of bet-hedging (i.e. obtaining increased fitness in stressful circumstances), evolutionary ecology can explain why parents have sometimes more offspring than they can afford (Kaplan, 1994; Lawson & Borgerhoff Mulder, 2016).

In economics, Classical theory, typically ascribed to Malthus, states that children are normal goods with no close substitutes and that the income elasticity for children is positive. However, this is at odds with the analysis of western countries passing through the demographic transition where the number of children declined over time (Fernihough, 2017). Becker (1960) generalized this model by proposing that parent's utility is defined by two dimensions: the quantity and quality of child-raising. He further assumes that the quantity of children can be fully controlled via contraception. Becker states that an increase in income due to the process of industrial revolution and the associated increase in the opportunity cost of child-raising results in two effects. The first effect is a positive income effect by the number of children. The second is the slightly higher negative substitution effect by the increased opportunity cost for child quality. This theory was later extended by Becker and Lewis (1973), Willis (1973), Becker and Tomes (1976), Becker, Murphy and Tamura (1990) and Moav (2005). Nonetheless, Galor (2012) reveals that, for countries who already passed the demographic transition period, fertility is not directly related with the income per capita across and within countries, refuting the income-related driving force that Becker ascribes to the trade-off. Additionally, Becker, Cinnirella and Woessmann (2010) state that Becker's models (1960; 1973) implicate an unexplained innate bias against child quantity (i.e. by assuming that the negative substitution effect is higher than the positive income effect).

Therefore, recent literature tries to explain the observed phenomenon with new theories. A very influential theory, the Unified Growth Theory, argues that technical progress related to the industrial revolution is the driving force resulting in an increase in income and demand for human capital (S. O. Becker et al., 2010; Cinnirella & Streb, 2017; Diebolt & Perrin, 2019; Galor, 2012; Galor & Weil, 1999, 2000). This results in reduced budget constraints and makes more resources available for the quantity and quality of children. Galor (2012), a proponent of the Unified Growth Theory, proposes the most recent model to clarify the possible existence of the quantity-quality trade-off. It starts from the proposition that the innate bias against child quantity originates from the decrease in gender wage gap (and hence the increase in women's labour force participation) resulting from the rise in demand for human capital. Consider a two-parents unitary household model where parents get utility[9] from both the number of (surviving) children (n) and the human capital of each child (h) (via increased returns on investment) (Eq. 2):

Eq. 2 $$U = (1 - \gamma) * \ln(c) + \gamma * [\ln(n) + \beta \ln(h)]$$

where 0<γ<1 and 0<β<1 represent the households' preferences towards child-raising and human capital respectively. Assume that child-raising is a task only done by women. This implies a cost defined as a fraction τ of women's unit-time endowment. This cost can be decomposed into the fraction of fixed time cost τ^q and the time cost τ^e per unit of education (*e*) (Eq. 3):

Eq. 3 $$\tau = \tau^q + \tau^e * e$$

The household's income consists of the men's and women's wage per unit of time (w^M and w^F respectively). This is divided over the opportunity cost of child-raising for the *n* children ($\tau * n * w^F$) and on consumption (*c*) (Eq. 4):

Eq. 4 $$\tau * n * w^F + c \le w^M + w^F$$

Eq. 4 reveals the innate bias against child quantity. If the women's relative wage increases, the opportunity cost of child-raising for the n children ($\tau * n * w^F$) increases more than the household income ($w^M + w^F$). Therefore, the negative substitution effect reflecting the increased opportunity cost of child-raising dominates over the positive income effect reflecting the increase in quantity of children. This causes a pressure to decrease the quantity of children. The optimal solution to this problem implies that (Eq.5):

Eq. 5 $$n = \frac{\gamma}{\tau^q + \tau^e * e}$$

which illustrates the inverse relationship (i.e. the trade-off) between quantity (*n*) and quality (*e*) of child-raising. Currently, the gender wage gap[10] in Ethiopia is estimated as 44% (Bank, 2019). If evidence would reveal the existence of a quantity-quality trade-off, this model shows that decreasing the gender wage gap could increase the negative substitution effect resulting in an accelerated decline in quantity.

2.4.1. Empirical evidence

Empirical evidence from the last ten years in developed and developing countries questions the existence of the quantity-quality trade-off (Clarke, 2018; Fernihough, 2017; Galor, 2012). However, some researchers investigate the causal effect in one of the two directions which is cautious since decisions towards quantity and quality are taken simultaneously (Cinnirella, 2019). In addition, the relationship is prone to endogeneity issues (Alidou & Verpoorten, 2019).

---

[9] The utility is assumed to be log-linear towards n, h and c in order to have a strictly monotonically increasing and quasi-concave function over [0, +∞[(Galor & Moav, 2002).

[10] Measured as the difference in average wages of men and women across both formal and informal sectors obtained from data from the Ethiopia Socioeconomic Survey from 2015-2016 (Bank, 2019).

Research that investigates the existence of the trade-off suggests that the trade-off is small or non-existent and can be non-linear (Clarke, 2018).

Focussing on SSA, empirical evidence is scarce. Recent literature reveals that some researchers find evidence for the trade-off (Bhalotra & Clarke, 2016; Bougma et al., 2015; Ito & Tanaka, 2017; Temel, 2013; Vogl, 2016) whereas others find no clear indication (Alidou & Verpoorten, 2019; Eloundou-Enyegue & Giroux, 2012; Kravdal et al., 2013). Nevertheless, it is interesting to analyse the existence of the trade-off in this study area for three reasons. First, most countries in SSA are assumed to be in stage two or three (i.e. they currently undergo the demographic transition). Second, the majority are confronted with budget constraints and child labour is still common (Alidou & Verpoorten, 2019). Third, SSA is still the region with low educational attainments and high TFRs. Only one recent study on the quantity-quality trade-off concerning education is found for Ethiopia which observes evidence for the existence of a trade-off (Gibson & Lawson, 2011).

This research examines if the level of fertility across households is inversely related with the level of education to add to the strand of literature. Our results can contribute to improve current family planning policies and interventions if the result would turn out that preferences for quality influence preferences for quantity. Besides, our contribution is that we assume a non-linear utility function, as we assume that when the quantity changes, this will change respondent's preference for quality, which was also argued by Mogstad and Wiswall (2009). Moreover, a mixture-amount model was used which is a useful model to study the quantity-quality trade-off due to the constraint that the total number of children attaining a certain educational level need to sum up to the total number of children. Nevertheless, this model is still quite uncommon in choice experiments, despite the work of some authors (Goos & Hamidouche, 2019; Khademi & Timmermans, 2012; Pradhan et al., 2017; Raghavarao & Wiley, 2009; Ruseckaite et al., 2017).

2.5. Gender

2.5.1. Theoretical evidence

Both the preference for investment in human capital and quantity of child-raising can be gender biased. First, concerning human capital, the gender gap in school enrolment rates in SSA is the highest in the world (United Nations, 2019a). The expected years of schooling in SSA is 9.8 years, with a gender gap of 9.3 years for girls and 10.3 years for boys. For Ethiopia, the expected years of schooling is lower, namely 8.7 years, with a gender gap of 8.3 years for girls and 9.1 years for boys. Second, concerning the preference for the number of children, the sex ratio at birth for SSA is 1.039 (meaning that for every girl born, there are 1.039 boys born) (place 152 out of 200 countries) and for Ethiopia this is 1.056 (place 148 out of 200 countries) whereas it occurs more often in South East Asia with China on place 1 with a sex ratio of 1.130 (The World Bank, 2020).

Theoretical literature concerning the gender issue in fertility is scarce. Eq. 2, 3, 4 and 5 can easily be extended by combining the model of Galor (2012) and de la Croix and Perrin (2016) with the addition of distinct utilities for boys (*b*) and girls (*g*) where the ratio of boys to girls is assumed to be one ($n_b = n_g = 0.5*n$) and the women's unit-time endowment for boys and girls also sum up to one ($\tau_b + \tau_b = 1$) (Eq. 6, 7, 8 and 9):

Eq. 6 $$U = (1-\gamma) * \ln(c) + \gamma * [\ln(n_b) + \ln(n_g) + \beta \ln(h_b) + \beta \ln(h_g)]$$

Eq. 7 $$\tau_b = \tau_b^q + \tau_b^e * e_b \text{ and } \tau_g = \tau_g^q + \tau_g^e * e_g$$

Eq. 8 $$\tau_b * n_b * w^F + \tau_g * n_g * w^F + c \leq w^F + w^M$$

Eq. 9 $$n = n_b + n_g = \frac{\gamma}{\tau_b + \tau_g} = \frac{\gamma}{\tau_b^q + \tau_b^e * e_b + \tau_g^q + \tau_g^e * e_g}$$

This shows the shrinking gender education gap (e_g/e_b) if either the fertility (*n*) or the price of teaching ($\tau^e * w^F$) decreases.

2.5.2. Empirical evidence

Recent empirical evidence concerning the gender gap in education in SSA reveals that there is evidence of sex preference in education (Bérenger & Verdier-Chouchane, 2016; Kazeem et al., 2010; Kuépié et al., 2015; Taş et al., 2014; Vimefall et al., 2017). In addition, recent studies in Ethiopia also conclude that there is sex preference in education (Mani et al., 2013; Tesfu & Gurmu, 2013).

The gender gap in preference for sons or daughters is extensively studied empirically in South (East) Asia. However, research in SSA is limited. Recent studies show that there is mixed evidence for the occurrence of a gender bias in SSA (Basu & De Jong, 2010; Bingenheimer & Raudenbush, 2004; Chao et al., 2019; Eliason et al., 2018; Flato, 2018; Fuse, 2010; Rossi & Rouanet, 2015). Additionally, husbands often have a higher preference for boys compared to wives (Root & Johnson-Hanks, 2016). Further, studies in Ethiopia also find mixed evidence (Basu & De Jong, 2010; Berlie & Alamerew, 2018; Fuse, 2010; Mekonnen & Worku, 2011; Rossi & Rouanet, 2015).

Therefore, it was chosen to add the gender aspect into our research. To our knowledge, recent studies in SSA focussing on the quantity-quality trade-off do not include this aspect.

2.6. Background section: Ethiopia

2.6.1. Ethiopia's demographic profile

Ethiopia is the 12th most populated country in the world and the 2nd in SSA (United Nations, 2019b). The population trends reveal a transition in the end of the second stage or beginning of the third stage of the demographic transition model in which the crude death rate in Ethiopia has decreased by 70% towards 6.7 in the past 50 years but the TFR declined more slowly with 37% towards 4.4 over the last 50 years (Figure 2-3) (The World Bank, 2020). This decrease is much steeper compared to SSA (a 59% decline in the crude death rate and a 29% decrease in the TFR over the last 50 years). The pattern of the recent steep decline in the TFR is also observed in other East-African countries compared to West- and Central Africa (Towriss & Timaeus, 2017). This makes Ethiopia an interesting case to study which factors explain the current lower TFR.

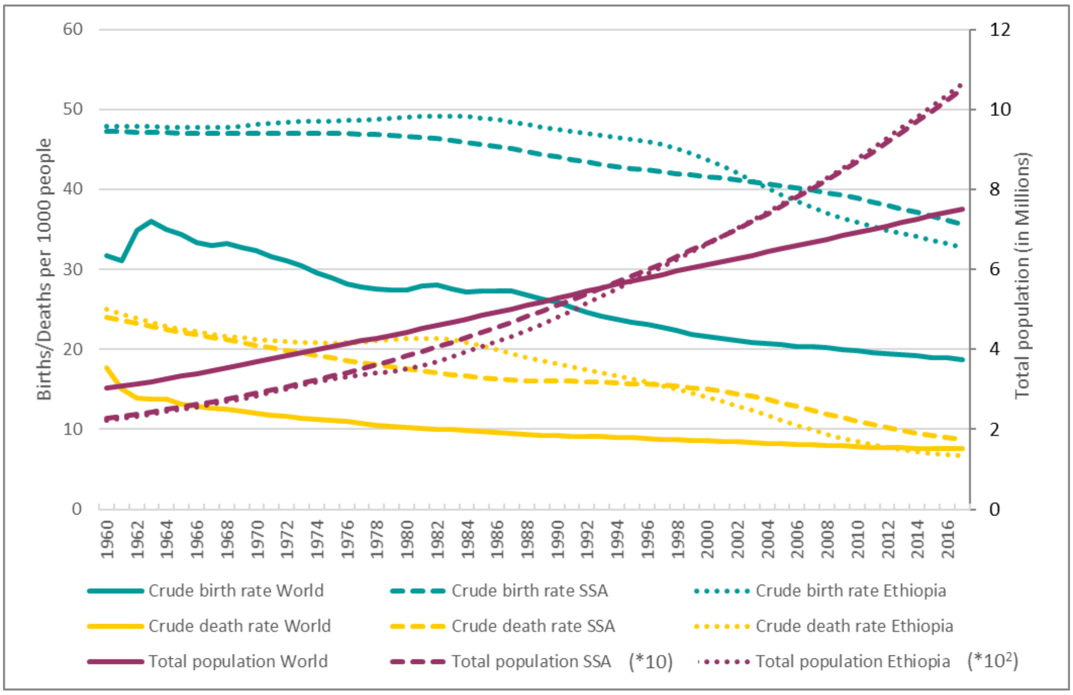


Figure 2-3: Evolution of the crude birth rate, crude death rate and the total population for the World, SSA and Ethiopia over the period 1960-2017 (The World Bank, 2018). The crude birth rate was chosen (although the TFR would be a better indicator as described in 2.1) to have an equivalent measure for the crude death rate.

The TFR is more than double in rural areas compared to urban areas (5.2 versus 2.3). This is due to the fact that Ethiopia is predominantly a rural society, with 79% living in densely populated rural settlements, which is much higher than the average in SSA of 60% (The World Bank, 2020). Moreover, Ethiopia is mainly a young society with 40% of the population under the age of 15 and only 4% above the age of 65 (Figure 2-4). This will probably result in a demographic dividend[11] followed by a proportional economic growth (Gerland et al., 2014). Further, extreme

[11] i.e. when the working-age population becomes larger than the non-working-age population who is dependent on the former (Gerland et al., 2014).

poverty[12] decreased with one percentage point between 2000 and 2015 (i.e. 30.8% of the population). In addition, the GINI[13] index is 39.1 (place 112 of 171), the Human Development Index[14] 0.463 (place 173 of 189) and the Gender Inequality Index[15] 0.502 (place 121 of 148) (The World Bank, 2020; United Nations, 2019a).

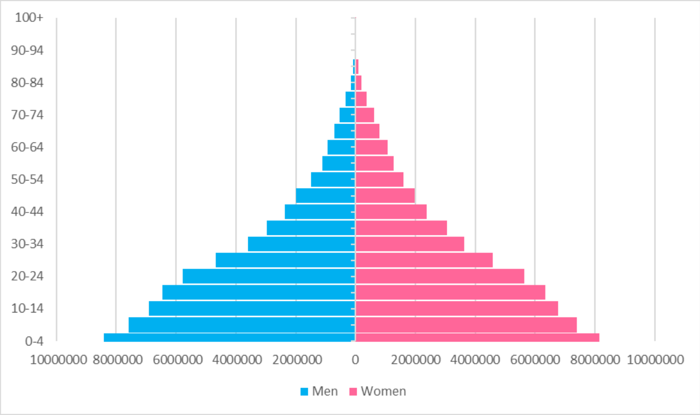


Figure 2-4: Population of Ethiopia by age groups for men (left) and women (right) respectively (United Nations, 2019b)

2.6.2. Ethiopia's family planning policies and programs

The main factors responsible for the progressive declines in fertility during the last 50 years include donor support, political will, non-governmental organisations (NGOs), public-private partnerships and the government's establishment of a network of Health Extension Workers (Olson & Piller, 2013). The NGO Family Guidance Association of Ethiopia, established in 1966, was the pioneer of the modern family planning services in Ethiopia (Olson & Piller, 2013). Since 1980, the Ministry of Health started to expand the family planning services, but the government lacked the capacity next to a lack of interest in the early 1990s, because Ethiopia's focus was on reconstruction after several years of war. The situation changed in 1993, when the government approved their first national population policy (Ringheim et al., 2009). At that time, the contraceptive prevalence was lower than in any other eastern and southern African country except Eritrea at that time (Central Statistical Agency & ICF International, 2011). The 1993 family planning policy increased the availability of contraceptives[16] from 4% in 1990 to 41% in 2017 with the support from NGOs (Central Statistical Agency & ICF International, 2019;

[12] i.e. living below the poverty line of 1.9$/day (2011 PPP) (The World Bank, 2020).
[13] Measures the deviation from a perfectly equal income distribution among households (corresponding with a value of 0) whereas a value of 100 corresponds with a perfectly unequal income distribution (The World Bank, 2020).
[14] Measuring the development of a country in terms of people and their capabilities along three dimensions: access to knowledge, a long and healthy life and a decent standard of living (The World Bank, 2020).
[15] Measuring the inequality between men and women along three dimensions: empowerment, reproductive health and the labour market (United Nations, 2019a).
[16] Measured as the percentage of women or their sexual partner are using any form of contraception (traditional as well as modern methods). Data is mainly obtained from household surveys where unmarried women are often excluded which can bias the estimates (The World Bank, 2020).

Ministry of Health [Ethiopia], 2010). This resulted in a decline in the TFR[17] from 7 in 1990 towards 4.4 in 2017 (The World Bank, 2020). In the following years, government support increased and different programs were developed (Olson & Piller, 2013). The annual per capita expenditure on health increased from $7 in 2005 to $28 in 2016, although this is much less than the World Health Organization's recommended $86 per capita (McIntyre, 2014; The World Bank, 2020).

The current policy is twofold. The first policy is the Costed Implementation Plan for Family Planning in Ethiopia (2016–2020) which was developed in line with the global strategy called Family planning 2020 (hosted by the United Nations) and in line with the national Health Sector Transformation Plan (2016–2020) (Ministry of Health [Ethiopia], 2016a). The policy focusses on increasing the contraceptive prevalence rate among youngsters and people living in low-resource settings and improving access to family planning information and services. The second policy is the National Adolescent and Youth Health Strategy (2016-2020) focussing on reducing adolescent pregnancy, preventable mortality, morbidity and suicide rate and reducing HIV incidence (Ministry of Health [Ethiopia], 2016b). Both policies were developed in line with the global strategy for Women's, Children's and Adolescents' Health (2016-2030) hosted by the World Health Organization. Currently, knowledge about contraception[18] is high (99%) with a domination of the method mix by injectables due to donor support (Central Statistical Agency & ICF International, 2016). However, there is a low contraceptive prevalence rate[19] (40%) (The World Bank, 2020) and the unmet need for family planning[20] in Ethiopia is estimated as 22% of which 13% wants to space births and 9% wants to limit births (Central Statistical Agency & ICF International, 2016). From 2000 towards 2016, government spending[21] on health decreased from 7.0% (4.4% of GDP) to 6.0% (4.0% of GDP) (The World Bank, 2020).

[17] Since 2012, birth registration is obligatory (UNICEF, 2019). However, currently only 2.7% completes birth registration resulting in the fact that fertility trends strongly depend on survey data (The World Bank, 2020).
[18] Measured as the percentage of men and women aging 15-49 year knowing at least one modern or traditional contraceptive method (Central Statistical Agency & ICF International, 2016).
[19] Measured as the percentage of men and women married or in union between 15 and 49 years old practicing any form of contraception (The World Bank, 2020).
[20] Measured as "the proportion of women who (i) are not pregnant and not postpartum amenorrhoeic and are considered fecund and want to postpone their next birth for 2 or more years or stop childbearing altogether but are not using a contraceptive method, or (ii) have a mistimed or unwanted current pregnancy, or (iii) are postpartum amenorrhoeic and their last birth in the last 2 years was mistimed or unwanted" (Central Statistical Agency & ICF International, 2016, p. 141).
[21] Measured as the percentage as total government expenditure (local, regional and central including transfers from international sources) (The World Bank, 2020).

2.6.3. Ethiopia's educational system

The educational system in Ethiopia is visualised in Figure 2-5. It starts at the age of seven with primary school, although pre-primary education exists in some regions (Education Policy Data Center, 2019; Trines, 2018). Primary school consists of two cycles (grade 1-4 and 5-8), or children can attend Alternative Basic Education (ABE). ABE is a schooling system for the first cycle of primary school in rural areas with a flexible structure to address the low attendance in these areas. After ABE, children can attend the second cycle of primary. Following primary school, there is a regional Ethiopian School Leaving Certificate exam. Passed children can continue secondary school which again consists of two cycles: high school (grade 9-10) and preparatory school (grade 11-12). Following high school, there is a second examination (the Ethiopian General Secondary Education Certificate Examination) whereby successful students can enter preparatory school (grade 11-12) or choose to follow technical and vocational education and training (TVET) institutions (consisting of 5 levels). TVET is concentrated in urban areas and provides a second path to enter university (after fulfilling two years of employment). Again, after preparatory school or eventually TVET, students can participate in the national exams (the Ethiopian Higher Education Entrance Examination) to enter university (3-5 years programme) or opt for TVET.

Primary and secondary is free in public schools, although paid public schools constitute 7% of the schools (mainly situated in Addis Ababa). In addition, TVET institutions are both private and public, but 75% are enrolled in private institutions which charge fees. Concerning universities, both public and private universities exist. Further, the teacher to child ratio in primary education is 1:55 (The World Bank, 2020). Although education is compulsory until the age of 16, only 54% of the enrolled students currently completes[22] primary education with a net enrolment rate[23] for primary school of 85% (The World Bank, 2020). Moreover, only 29% completes secondary school with a net enrolment rate of 31%. In addition, the net enrolment is 8% for tertiary school (no data for the completion rate is found).

[22] Measured as the ratio of the number of entrants (regardless of their age) into a program to the population of the official school age (The World Bank, 2020).

[23] Measured as the ratio of children at the official school age enrolled in the program to the population of the official school age based on annual school surveys (The World Bank, 2020).

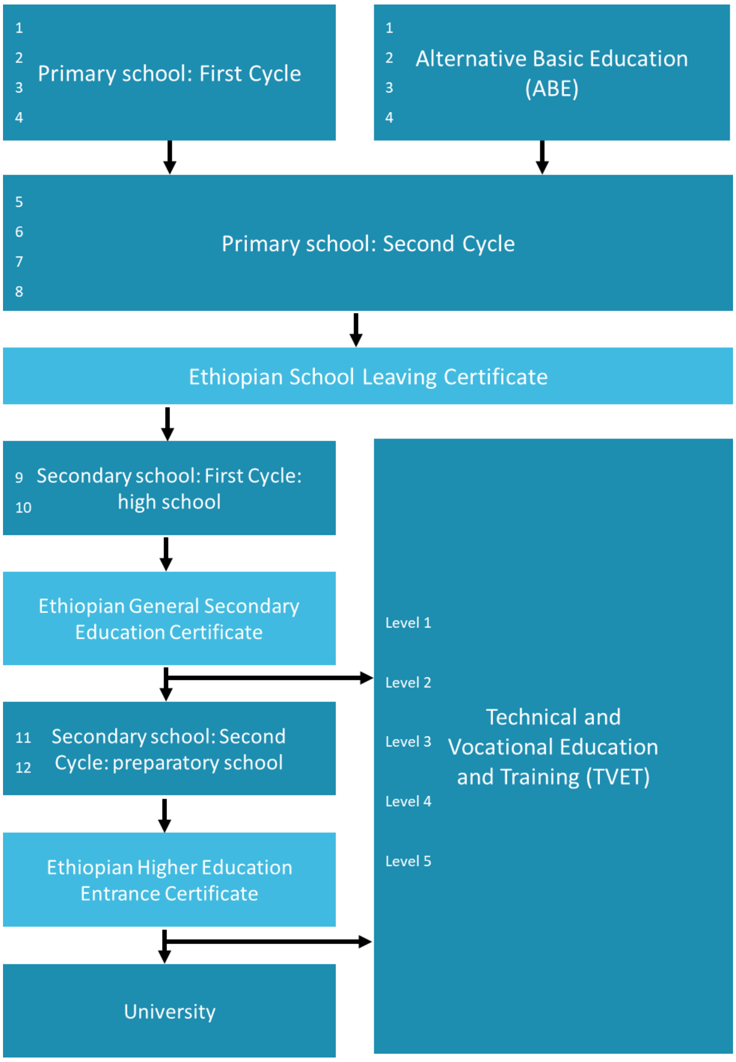


Figure 2-5: Visualization of the educational system in Ethiopia

In 1994, the New Education and Training Policy was put into action which gave education a developmental priority on the agenda. The main aim was to improve and to provide a more equitable access to education. Since 1997, the Education Sector Development three/five-year Program was adopted of which now round five is into force. The main goal is capacity development for improved management, quality, enrolment rates and efficiency (Federal Ministry of Education, 2015). Furthermore, it includes specific targets to achieve gender parity in the education and training sector. From 2000 to 2015, government spending[24] on education decreased from 12.4% (2.1% of GDP) to 27.1% (4.7% of GDP), but support by donors is still a large source of funding for primary schooling (Oumer, 2009; The World Bank, 2020).

[24] Measured as the percentage as total government expenditure (local, regional and central including transfers from international sources) (The World Bank, 2020).

3. Methodology

3.1. Study area

Administratively, Ethiopia is divided into nine regional states (also called kililochs) based on ethnic territoriality, each consisting of different zones and subsequently divided into districts (or woreda's) covering a total of around 15 000 municipalities (or kebeles) (Central Statistical Agency, 2012). To achieve the study objectives, a choice experiment was implemented in two zones: the Gamo Gofa Zone and the Segen People's Zone of the Southern Nations, Nationalities, and People's Region (SNNPR) in Ethiopia (Figure 3-1). These two zones were selected since this thesis is part of the current VLIR-UOS financed Interuniversity Cooperation Program in the same region, called "Living with uncertainty: Analysing rural livelihoods and rethinking sustainability in the South Ethiopian Rift Valley". These two administrative zones cover the third and fifth largest areas of the 22 zones of the SNNPR respectively. Within these zones, six districts were selected, namely Arba Minch Zuriya, Bonke, Chencha, Merab Abaya, Konso and Derashe. The research area has a tropical semi-arid to (sub-)humid climate with an annual temperature of 10.1-27.5 °C, an annual precipitation of 801-1800 mm and altitudinal variations of 501-3500 m (The Ethiopian National Meteorological Services Agency, 2019).

The study area comprises a population of 965,722 inhabitants (Table 3-1), accounting for around 1% of the total population of Ethiopia and has a mean population density of 180.28 inhabitants/km² (Central Statistical Agency, 2012; United Nations, 2019b). The zones are inhabited by 85 ethnic groups and various religions[25]. The most common ethnicities in the Gamo Gofa Zone are the Gamo (64.61%) and the Gofa (22.08%) (Central Statistical Agency, 2012). In the Segen People's zone, mainly the Konso are most present (30.40%), followed by the Kore (21.48%), the Oromo (13.47%) and the Debase/Gewada (9.15%). Concerning religion, 52.22% are Protestant in the Gamo Gofa Zone, while 32.69% are Orthodox and 12.66% practice a traditional religion (Central Statistical Agency, 2007b, 2007a). In the Segen People's zone, 56.14% is Protestant, 20.40% Orthodox, 16.08% Traditional and 5.92% Catholic. Both the Gamo Gofa and Segen People's Zones show a decline in TFR in the last decade from 5.6 and 6.1 in 2000 respectively to 4.9 and 5.6 in 2011 (Teklu et al., 2013) and to 4.4 in 2014 for the SNNPR in general (Central Statistical Agency & ICF International, 2014).

[25] Neither in the census of 2007 or 2012 specifications of the category "other" religions was provided. Therefore the number of religions is not documented.

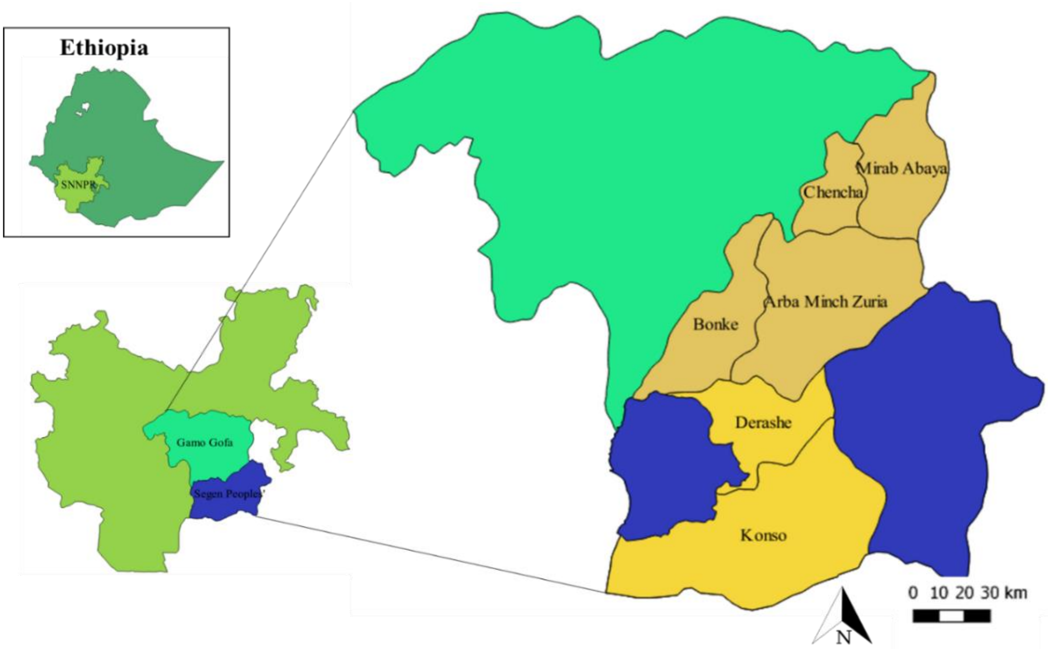


Figure 3-1: Map of the study area in the Goma Gofa Zone and Segen People's Zone, Southern Nations, Nationalities, and People's Region (SNNPR) in Ethiopia

Table 3-1: Demographic statistics of the study area (Central Statistical Agency, 2007b, 2012; United Nations, 2019b)

Regional states *(kililoch)*	District *(woreda)*	Area (km²)	Population			Mean population density per km²	% of the population living in rural areas
			Total	Men	Women		
Goma Gofa zone							
	Arba Minch	967.69	187,811	93,829	93,982	194.08	68.72
	Bonke	792.37	182,946	90,967	91,979	230.88	96.01
	Chencha	373.52	130,309	59,958	70,351	348.87	88.09
	Mirab Abaya	630.76	86,811	43,362	43,449	137.63	92.23
Segen Peoples zone							
	Konso	2 273.79	235,087	113,412	121,675	103.39	96.00
	Derashe	1 487.38	142,758	70,111	72,473	95.98	90.76

Different valuation methods can be used to answer the research questions. An economic valuation was chosen to reveal quantitative outcomes (Kjaer, 2005). Further, a mixed method approach was used to increase the reliability (Powe et al., 2005). Since the quantity and quality of child-raising is a non-monetary value, first a non-monetary deliberative and participatory approach, namely a focus group discussion was performed. This reveals the respondent's self-expressed preferences. Additionally, stated preferences were analysed, since fertility decisions can be considered as a hypothetical market with a demand and supply side (Perman et al., 2011). Two techniques are available: a DCE and contingent valuation. Since the contingent valuation method is not able to reveal the trade-off respondents make between different characteristics of options, it was more appropriate here to use a DCE (Kjaer, 2005).

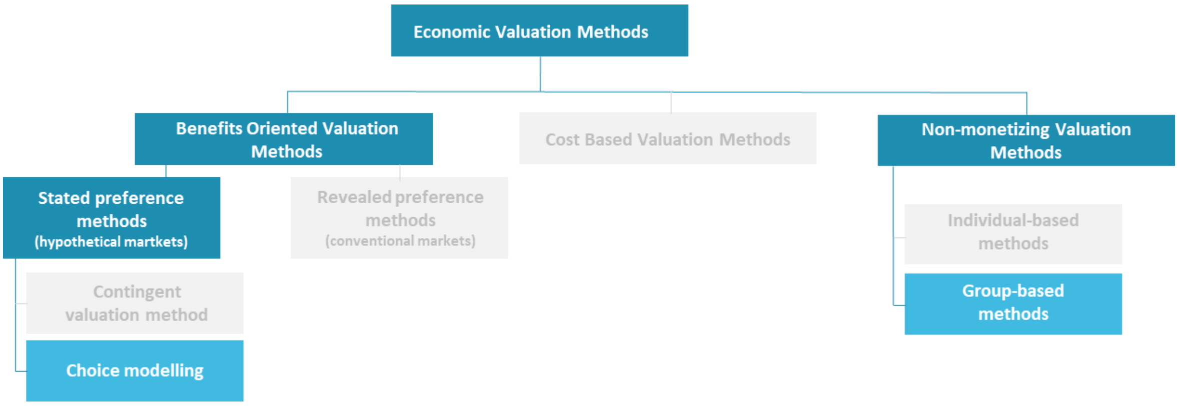


Figure 3-2: Schematic visualization of different economic valuation methods (adapted from Forest Europe (2019))

3.2. Focus group discussions

A focus group discussion is defined as a well-planned discussion to gather a lot of information via analytic induction on a group-level in a short time frame (Merton & Kendall, 1946; Mortelmans, 2013; Znaniecki, 1934). It is frequently used in combination with quantitative survey research as a preliminary investigation to help establishing the quantitative survey or as a follow-up investigation in order to deepen the results obtained from the quantitative analysis (Morgan, 1996). In focus group discussions, the interaction between the group members is important, because discussions can reveal information that is only the result of group interactions. This leads to gathering more information than just the answers on the respondent's questions in case of individual interviews (Barbour & Kitzinger, 1999; Bloor et al., 2001). Besides, the process through which the ideas and opinions are created can be observed, next to what the respondents really mean with their answers (Greenbaum, 1998).

In this research, focus group discussions were held before the DCE for three reasons. First, this was done because of language and cultural differences (Mangham et al., 2009). Second, this was conducted to minimize specification errors of the DCE and follow-up questions which can occur during the conceptualization and the operationalization of the research questions (de Leeuw et al., 2008). Third, this was performed to reduce measurement errors (i.e. the difference between the true and the observed value (de Leeuw et al., 2008)). The flowchart of the implementation of the focus group discussions is shown in Figure 3-3.

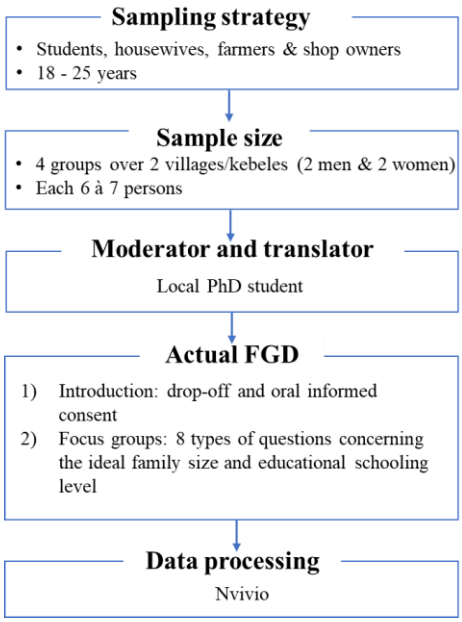


Figure 3-3: Visualization of the workflow of the implementation of the focus group discussions

The sampling strategy was consistent with a targeted theoretical or constructivist sampling approach which believes that it is more meaningful to sample units based on their specific characteristics to include different perspectives on the subject (Roose & Meuleman, 2014). This view is opposed to the naturalistic approach which states that an objective reality can be created by sampling a random sample. More specifically, respondents were selected based on discriminatory sampling in order to maximize socio-economic differences between the respondent groups and to minimize the differences within the groups (Merton & Kendall, 1946; Mortelmans, 2013; Roose & Meuleman, 2014; Van Hove et al., 2014). In homogeneous groups, respondents feel more comfortable and are more willing to provide relevant information. This was combined with a snowball sampling approach in which respondents were asked if they knew other potential respondents from a specific population (Roose & Meuleman, 2014; Van Hove et al., 2014). Finally, the following demographic categories were included: students, house wives, farmers and shop owners all between the age of 18 and 25. The descriptive sampling matrix can be found in Table 8-1 in Appendix.

The sample size was determined by saturation of the gathered information (i.e. theoretical saturation) (Fern, 2001; Morgan, 1998). In practice, each focus group was analysed afterwards based on the ‘constant comparative method’ (Glaser & Strauss, 1967). The concepts found from the analysis of the data were constantly compared and adjusted with the new information until saturation was reached. This resulted in four focus groups of six to seven people who came together for approximately one hour and a half spread over two villages. There were two focus group discussions with women and two with men. Potential participants were personally invited by the village leader two days before the interviews. This was done by asking a list of potential candidates to the administration centre of the villages. Moreover, the day itself, every respondent was called back as a reminder.

Krueger (1988) states that the moderator at best resembles the respondents to reduce the effect of power resulting from a dominant position as moderator. Therefore, the focus group moderator was a local PhD student, who was able to translate. Next to personal characteristics, situational elements and training are important because the moderator has to fulfil different roles: as interviewer, facilitator, leader, psychologist and time guard (Bloor et al., 2001).

Before implementing the actual focus group discussions, the respondents were asked to fill in the drop-off (i.e. a short, closed questionnaire to obtain the same background information about every respondent) (Figure 8-1 in Appendix). Some questions include the names of the partner and children to facilitate the discussion afterwards (Mortelmans, 2013). In addition, they were asked to take note and give approval to the informed consent (Figure 8-2 in Appendix). The informed consent was done oral, seen that the adult literacy rate is only 39% (Central Statistical Agency & ICF International, 2016). The questions asked during the focus group discussions concerned the following topics: ideal family size, the ideal educational level and gender composition (Table 8-2 in Appendix).

During the focus group discussions, the interviews were transcribed at the moment itself with the help of the translator. Afterwards, this was analysed using Nvivo 12 Pro (QSR International, 2018). The interviews were all anonymized to guarantee the privacy of the respondents.

3.3. Discrete Choice Experiment

3.3.1. Theoretical framework

A DCE is a method used to reveal people’s stated preferences, in this case fertility preferences, in the form of a survey (Louviere et al., 2010). It is commonly used in health economics, agricultural and food economics, and environmental and resource economics. It is based on the random utility theory proposed by Thurstone (1927) and further elaborated by McFadden (1974) and Manski (1977), and on the characteristics theory of value behaviour (Lancaster, 1966).

In practice, distinct choice cards are given repeatedly to the respondents of which an example is shown in Figure 3-4 (Kjaer, 2005). Each card consists of two or more mutually exclusive alternatives (i.e. options) of which respondents are asked to choose their most preferred one. Alternatives are defined by a set of features (i.e. attributes) which possess different categorical or continuous levels (i.e. the number of children in Figure 3-4). Therefore, the respondents have to make a trade-off between the attribute levels. Next to the alternatives, an opt-out option is often added which equals the status-quo to avoid that respondents make forced choices (Hoyos, 2010). Analysis of this data can reveal which attributes and levels contribute most in choosing a certain alternative. This can be done using different discrete choice models.

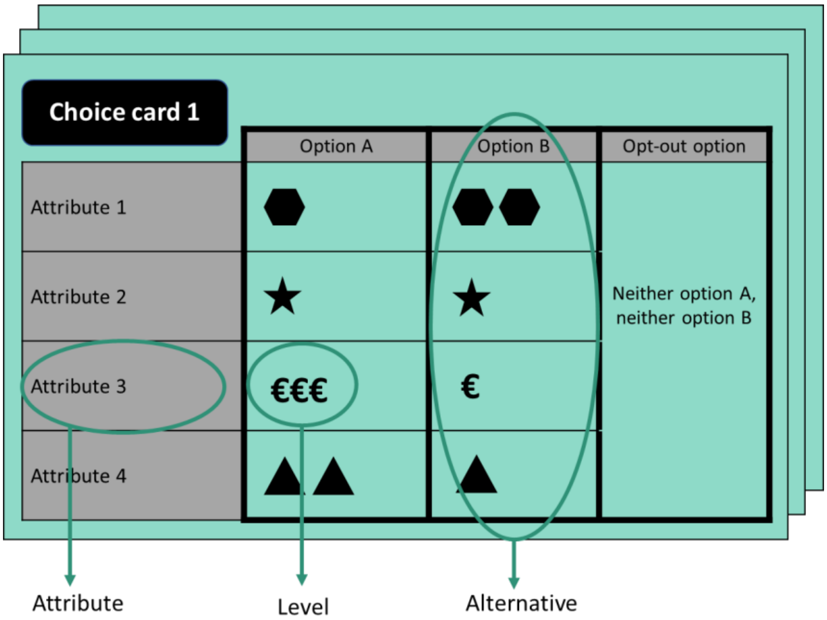


Figure 3-4: Example of a choice card

Random utility theory states that every person has a latent 'utility' for each alternative which we cannot directly observe (Louviere et al., 2010). This utility depends on the individual's socio-economic characteristics and the attributes. It can be split up in two elements: a systematic or observable and a random or unobservable component (Eq. 10):

Eq. 10 $$U_{ni} = V_{ni} + \varepsilon_{ni}$$

where U_{ni} is the latent utility of individual *n* for alternative *i*. The systematic element (*V*) can be approximated by a function of explanatory variables capturing the attributes (which comprises the differences in alternatives) and the covariates (which covers the differences between individuals). The random component (ε) includes the random variation in the choice-making process. Because of the random component, utilities are stochastic and can be estimated using different probabilistic discrete choice models based on (Eq.11):

Eq. 11 $$P(i|C_n) = P\left[V_{ni} + \varepsilon_{ni} \geq V_{nj} + \varepsilon_{nj}\right] = P\left[V_{ni} - V_{nj} \geq \varepsilon_{ni} - \varepsilon_{nj}\right] \forall\, i \neq j;\ i,j \in Cn$$

which states that the probability that individual *n* selects option *i* from the choice set C_n is equal to the probability that the systematic and random component of the utility of option *i* is larger than the utility of option *j* from the same choice set based on utility maximizing behaviour of the respondents. This expression shows that it are differences in the random and stochastic components that are important (Lancsar et al., 2017). A more general expression involving two or more competing options in choice set C_n can be written as follows (Eq. 12):

Eq. 12 $$P(i|C_n) = P[(V_{ni} + \varepsilon_{ni}) \geq Max(V_{nj} + \varepsilon_{nj})] \; \forall \; i \neq j; \; i, j \in Cn$$

Only certain specifications of the distribution of ε lead to closed-form expressions allowing the parameters to be estimated. These discrete choice models are explained under 3.3.4.1.

In practice, first, the experiment needed to be designed whereby the characteristics of the choice options were systematically changed corresponding to an experimental design (Goos & Hamidouche, 2019). Next, the DCE was implemented in the field. Afterwards, the data was analysed using different discrete choice models. These three steps will be discussed in the following paragraphs.

3.3.2. Design

Figure 3-5 gives an overview of the design of the DCE. First, the research question was decomposed into relevant attributes. Then, the levels of the attributes were used to design the experiment via optimal design of experiments techniques. Finally, the choice cards were made based on this design. The following paragraphs describe the procedure in detail.

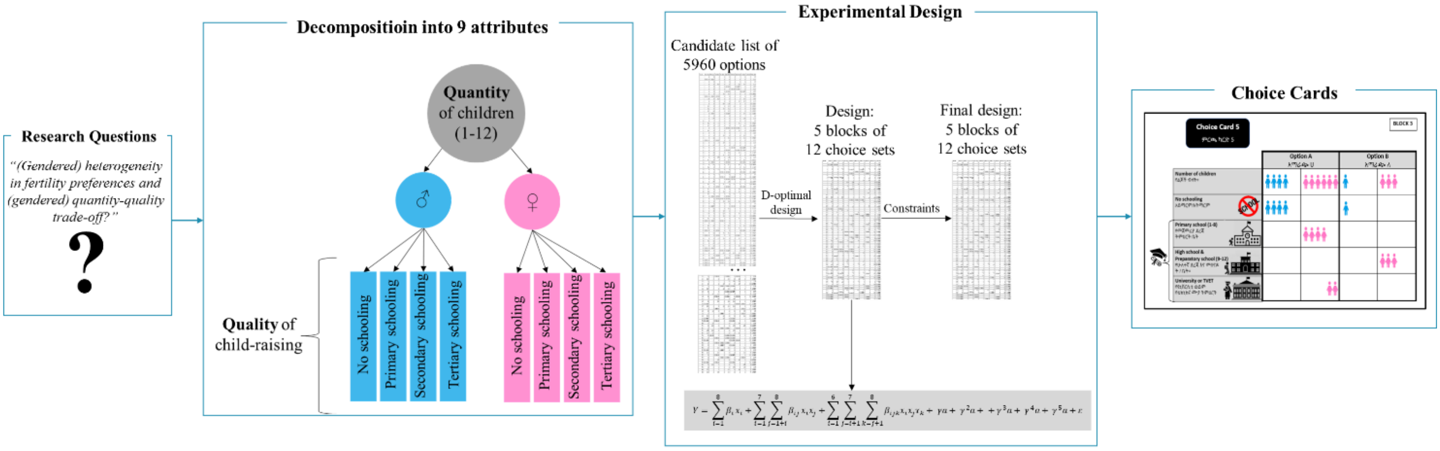


Figure 3-5: Visualization of the workflow to design the DCE. This consists of a decomposition of the research question into the relevant attributes (i.e. features). Then, the levels of these attributes were selected based on an experimental design. Afterwards, the choice cards were made based on this design. An enlarged version of the choice card can be found in Figure 3-8.

3.3.2.1. Selection of the attributes

The selection of the attributes was based on the existing literature and in cooperation with the thesis supervisors resulting in a focus on quantity, quality and gender of child-raising. To decide how many attributes should be included, a trade-off was made between the complexity of the task for the respondents and the omitted variable bias (Hoyos, 2010). Consequently, it was chosen to include the attribute *number of children* (as the sum of the number of boys and girls) and the division of this total number of children over eight categories: boys/girls not receiving any schooling and boys/girls attaining a certificate at the end of primary, secondary (including both high school and preparatory school) or tertiary (including both TVET and university) schooling respectively (Table 3-2). This allocation of the different educational levels can be seen as proportions of the total number of children which then sum up to one. So, the fertility preferences depend on the total amount and on the proportions of the number of boys and girls attaining each educational level. Modelling this design involves a special type of model, called a mixture-amount model.

The categorizations into four educational levels was based on the fact that TVET and university are more costly since primary and secondary school is for free while TVET and university are asking a school-fee. Moreover, TVET institutions are more centralized in the larger cities in the districts resulting in higher transportation and/or accommodation costs (The World Bank, 2020). Further, the choice to focus on boys/girls finishing a certain educational level was based on the high drop-out ratio during primary school (46%), secondary school (71%) (The World Bank, 2020).

The attributes consist of different discrete levels ranging from zero/one to twelve children including gender specification (Table 3-2). The maximum was set at 12 children for two reasons. First, the maximum number of children in a previous household survey in the study area was 16, occurring only in 2 of the 879 households (Feyisa, 2019). Second, the TFR in the study area is 4.9 in the Gamo Gofa Zone and 5.6 in the Segen People's Zone (Teklu et al., 2013). A trade-off was made in keeping the number of attribute levels as low as possible in order to reduce the number of possible combinations and to include enough intermediate options. Still, this resulted in 11^8*10^1 = 2,143,588,810 possible attribute level combinations (8 attributes concerning the educational level with each possessing 11 attribute levels and 1 attribute for the amount variable with 10 levels). However, not all combinations are feasible. Therefore, the constraint that the sum of the number of girls and boys distributed over the 4 educational levels has to sum up to 12 was added. This resulted in 162,425 possible combinations.

Table 3-2: Attributes and their levels in the mixture-amount model

Component of the model	Variable in the model	Attributes	Attribute levels
Amount component	a	total number of children (NChildren)	1, 2, 3, 4, 5, 6, 8, 9, 10, 12
Mixture component	x_1	number of boys who have not received any schooling (MNoPrimary)	0, 1, 2, 3, 4, 5, 6, 8, 9, 10, 12
	x_2	number of girls who have not received any schooling (FNoPrimary)	0, 1, 2, 3, 4, 5, 6, 8, 9, 10, 12
	x_3	number of boys who have completed primary schooling (MPrimary)	0, 1, 2, 3, 4, 5, 6, 8, 9, 10, 12
	x_4	number of girls who have completed primary schooling (FPrimary)	0, 1, 2, 3, 4, 5, 6, 8, 9, 10, 12
	x_5	number of boys who have completed high school & preparatory schooling (MSecondary)	0, 1, 2, 3, 4, 5, 6, 8, 9, 10, 12
	x_6	number of girls who have completed high school & preparatory schooling (FSecondary)	0, 1, 2, 3, 4, 5, 6, 8, 9, 10, 12
	x_7	number of boys who have completed university or TVET (MTertiary)	0, 1, 2, 3, 4, 5, 6, 8, 9, 10, 12
	x_8	number of girls who have completed university or TVET (FTertiary)	0, 1, 2, 3, 4, 5, 6, 8, 9, 10, 12

3.3.2.2. Mixture-amount modelling

The following paragraphs provide an explanation of the theoretical background about the mixture-amount modelling of the levels of the attributes.

Mixture models

A simple mixture model contains the unique feature of the mixture constraint whereby, for all the p ingredients, the proportions x_i of the i^{th} ingredient in the mixture sum up to one (Eq. 13) (Goos & Jones, 2011):

Eq. 13

$$\sum_{i=1}^{p} x_i = 1$$

An example of a mixture model is the quality of a cake which depends on the ingredients. Graphically, the mixture constraint for three ingredients is illustrated in Figure 3-6. The mixture constraint implies that the intercept is redundant seen that the mixture proportions sum up to one which results in linear dependence or perfect collinearity among the parameter columns in which case the information matrix used for the estimator is singular.

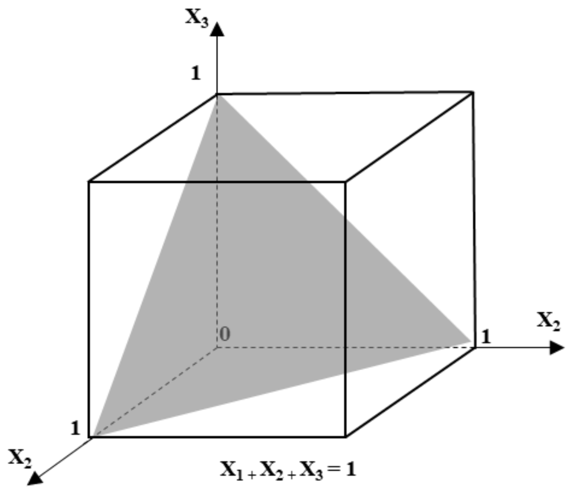


Figure 3-6: Experimental space for a mixture with three components (adapted from Goos and Jones (2011)*)*

For the chosen attributes stated above, there are 8 ingredients, namely the number of boys and girls divided over the 4 educational levels which sum up to a maximum of 12 (which were normalized to proportions which sum up to 1 in the experimental design). The mixture constraint implies that if one proportion changes, the other proportions will automatically follow. The result is that the proportions cannot be orthogonal so that the effects cannot be estimated independently (Aleksandrovs et al., 2015). This makes most of the usual significance tests useless since these assume an independent interpretation of the regression coefficients. Still, the models can be useful for predictions and optimizations.

Scheffé (1958) was the first to describe mixture models. He proposed different well-known polynomials to model data involving mixture experiments: a first order or linear (Eq. 14a), second order or quadratic (Eq. 14b), third order or special cubic (Eq. 14c) and fourth order or full cubic model (Eq. 14d):

Eq. 14a
$$Y = \sum_{i=1}^{p} \beta_i x_i + \varepsilon$$

Eq. 14b
$$Y = \sum_{i=1}^{p} \beta_i x_i + \sum_{i=1}^{p-1} \sum_{j=1+i}^{p} \beta_{ij} x_i x_j + \varepsilon$$

Eq. 14c
$$Y = \sum_{i=1}^{p} \beta_i x_i + \sum_{i=1}^{p-1} \sum_{j=1+i}^{p} \beta_{ij} x_i x_j + \sum_{i=1}^{p-2} \sum_{j=i+1}^{p-1} \sum_{k=j+1}^{p} \beta_{ijk} x_i x_j x_k + \varepsilon$$

Eq. 14d
$$Y = \sum_{i=1}^{p} \beta_i x_i + \sum_{i=1}^{p-1} \sum_{j=1+i}^{p} \beta_{ij} x_i x_j + \sum_{i=1}^{p-1} \sum_{j=1+i}^{p} \delta_{ij} x_i x_j (x_i - x_j) + \sum_{i=1}^{p-2} \sum_{j=i+1}^{p-1} \sum_{k=j+1}^{p} \beta_{ijk} x_i x_j x_k + \varepsilon$$

with Y the outcome variable, β the vector of parameter estimates, x the mixture components and ε the independent and identically distributed (I.I.D.) error term following a normal distribution with mean zero. The cross-product terms involve the possible interactions between the proportions of the mixture components. If the sign of the corresponding regression coefficient β is positive, there is a synergetic effect of the mixture components; otherwise the effect is antagonistic. These models do not contain a quadratic effect of the mixture components. The reason is that the quadratic effect is redundant because it is equal to a linear combination of the linear and two-factor interaction effects of the proportions composing the mixture shown in Eq.15:

Eq. 15
$$x_i^2 = x_i(1 - \sum_{\substack{j=1 \\ j \neq i}}^{p} x_j) = x_i - \sum_{\substack{j=1 \\ j \neq i}}^{p} x_i x_j$$

Later, other models were introduced: a model with inverse terms (Draper & St John, 1977), a model for inactive or additive components (N. G. Becker, 1968; N. S. Becker, 1978), a ratio model (John A. Cornell, 1981; Snee, 1973) and a log-ratio model (Aitchison & Bacon-Shone, 1984).

Mixture-amount models

The response can also depend on process variables of the mixture components, next to the mixture proportions. The previous example of the quality of a cake can depend not only on the ingredients but also on the baking time and temperature for example. For analysing this kind of experiments, the mixture models for the mixture proportions can be combined with a response surface model for the process variables, first introduced by Cornell (1971). A response surface model uses a second-order Taylor series. This results in a model with quadratic effects of the continuous factors to model the curvature between these factors and the response variable. An example is given in Eq. 16 (Goos & Jones, 2011). This requires a minimum of three levels for the continuous factors in order to model the curvature.

Eq. 16

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i \, x_i + \sum_{i=1}^{k-1} \sum_{j=1+i}^{k} \beta_{ij} \, x_i x_j + \sum_{i=1}^{k} \beta_{ii} \, x_i^2 + \varepsilon$$

A special type of a process variable is the total amount of the mixture components. This model is called a mixture-amount model in literature (Piepel & Cornell, 1985). An example is the application of fertilizer which depends on the fertilizer composition but also on the amount applied. Different models have been proposed to estimate this where the effect of the amount variable can be additive or multiplicative. Consider a second order Scheffé model for the mixture variables (Eq. 14b) and a response surface model for the amount variable (Eq. 16). For an additive amount effect of power n this can be written as (Eq. 17a):

Eq. 17a

$$Y = (\sum_{i=1}^{p} \beta_i \, x_i + \sum_{i=1}^{p-1} \sum_{j=1+i}^{p} \beta_{ij} \, x_i x_j) + \gamma a + \gamma a^2 + \cdots + \gamma a^n + \varepsilon$$

where *a* is the total amount belonging to the interval [1, 12] in this research. This is the case if one assumes that the mixing of the mixture components is the same at all levels of the process variables (Kowalski et al., 2000). If it is assumed that the process variables or *amount* depend in reality on the relative proportions of the mixture components, the interaction terms between both can be included in the model. For a multiplicative amount effect of power *n*, this becomes (Eq. 17b):

Eq. 17b

$$Y = (\sum_{i=1}^{p} \beta_i x_i + \sum_{i=1}^{p-1} \sum_{j=1+i}^{p} \beta_{ij} x_i x_j) + a(\sum_{i=1}^{p} \beta_i x_i + \sum_{i=1}^{p-1} \sum_{j=1+i}^{p} \beta_{ij} x_i x_j) + a^2(\sum_{i=1}^{p} \beta_i x_i + \sum_{i=1}^{p-1} \sum_{j=1+i}^{p} \beta_{ij} x_i x_j) + \cdots + a^n(\sum_{i=1}^{p} \beta_i x_i + \sum_{i=1}^{p-1} \sum_{j=1+i}^{p} \beta_{ij} x_i x_j) + \varepsilon$$

where a is again the total amount belonging to the interval [1, 12]. These models do not contain the main effect of the process variables because including process variables with two-factor interaction terms of the process variable and the mixture components results in (Eq. 18):

Eq. 18

$$z * \sum_{i=1}^{p} x_i = z$$

The number of parameters can be quite large if the number of mixture components increases. This might result in overfitting and the need of large datasets which is often practically impossible due to time and cost constraints. Therefore, some reduced models have been proposed such as the ones from Kowalski et al. (2000) or Prescott (2004). Further, more recently, other models than the Scheffé models have been proposed as listed by Khuri (2006) and Pal and Kumar (2012).

3.3.2.3. Experimental design

Mixture experiments are still quite uncommon in choice experiments, despite the work of some authors (Goos & Hamidouche, 2019; Khademi & Timmermans, 2012; Pradhan et al., 2017; Raghavarao & Wiley, 2009; Ruseckaite et al., 2017). In this section, the method used for obtaining an efficient experimental design for the choice experiment is explained. The experimental design is closely linked to model specification in the analysis step because the types of models that can be estimated are determined within the constraints and statistical properties of the experimental design (Lancsar et al., 2017; Zijlstra et al., 2019).

The choice cards were designed using the JMP Pro 14 (SAS Institute Inc., 2018) and SAS 9.4 software (SAS Institute Inc., 2013). Seen that the JMP software was not able to design a mixture-amount experiment, the design involved two steps (Figure 3-5). **First**, a candidate list for the mixture proportions was made in JMP. This was created via a simplex lattice design (i.e. an angular grid) at proportions of 1/n (with n the different attribute levels considered above) that gave rise to integer numbers of children (ranging from 1 to 12). An example of such a simplex lattice design visualised in a triangular grid for three of the attributes in this experiment is given in Figure 3-7. The candidate list of the mixtures was repeated for different combination of numbers of boys and girls. This resulted in a candidate list of 5960 possible options.

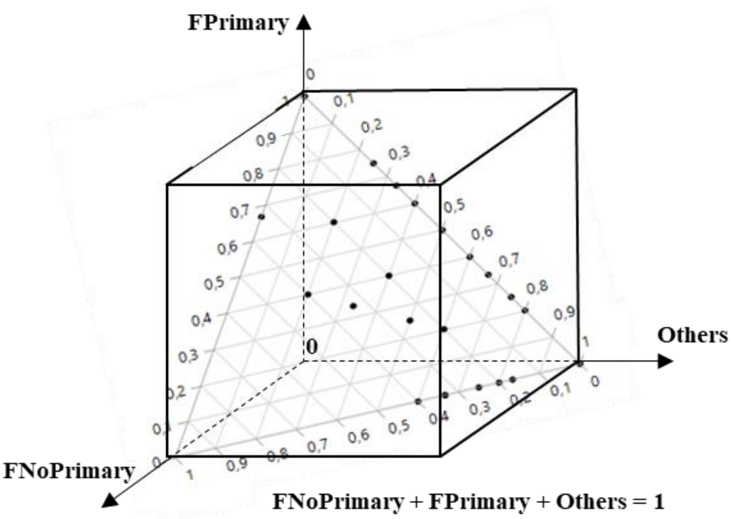


Figure 3-7: Visualization of a part of the final experimental design in this research (based on the simplex lattice design) in a triangular grid (adapted from Goos and Jones (2011)). The black dots indicate the settings of the mixture components. Since our design included eight mixture components, the full visualization would require an eight-dimensional plot.

Second, the choice sets with the highest information content needed to be selected from the candidate set. This was done via a fractional factorial design, namely a D-optimal design using a mixture-amount model in SAS 9.4 for two reasons. First, a full factorial design consisting of all attribute level combinations was not feasible. Second, optimal designs are getting growing support compared to orthogonal designs since it requires less runs and no constraints on the design space (Goos & Jones, 2011). A D-optimal design searches for the highest information content by maximizing the information matrix of the parameter estimates in order to minimize the variance-covariance matrix of the parameters resulting in the most precise parameter estimates (Goos & Jones, 2011; Street et al., 2005).

However, it must be taken into account that in reality the family size is approximately normally distributed instead of uniformly distributed (The World Bank, 2020). Additionally, based on previous literature, it was hypothesized that respondents have a preference for a large number of children (G. S. Becker, 1960; Galor, 2012), mainly boys (Bongaarts & Casterline, 2013) and a high educational level (Galor, 2012). But SAS 9.4 cannot build a Bayesian design (i.e. a design in which the model parameters equal a value based on prior knowledge regarding the magnitude and sign (Kessels et al., 2008)) or an I-optimal design (i.e. a design which minimizes the average variance of prediction resulting in more data points in the middle range compared to a D-optimal design (Goos et al., 2016; Goos & Jones, 2011)). One remedy would have been to add a status quo option. However, this is an 'easy-out' option which could result in a lack of information seen that previous research revealed that parents have a lot of self-expressed reasons related to situational influences (e.g. the decision of God) rather than personal dispositional factors (Farina et al., 2001). Another reason is that this research is an ex ante study where it was assumed that respondents want to have children in the future. Another solution would be to conduct a small pilot study as proposed by Huber & Zwerina (1996). Due to time constraints, this was not a feasible option.

Therefore, an alternative solution was implemented using a D-optimal design with a rather complex model in order to make it possible to add more intermediate attribute levels in the design. The model used for creating the experimental design was a third order Scheffé model for the eight mixture components combined with an additive response surface model of order five for the amount effect (Eq. 19):

Eq. 19 $$Y = \sum_{i=1}^{8} \beta_i x_i + \sum_{i=1}^{7} \sum_{j=1+i}^{8} \beta_{ij} x_i x_j + \sum_{i=1}^{6} \sum_{j=i+1}^{7} \sum_{k=j+1}^{8} \beta_{ijk} x_i x_j x_k + \gamma a + \gamma^2 a + + \gamma^3 a + \gamma^4 a + \gamma^5 a + \varepsilon$$

This design is a compromise between statistical efficiency for a meaningful regression model and the requirement that the profiles offered are realistic, varied, and contain no or few dominant profiles within a single set of options. The model contains five features. First, the third order Scheffé model allowed to take into account potential curvature in the model. Second, this higher-order model allowed to estimate any simpler model in the analysis step after obtaining the data. Third, the response surface model of order five was added not because it was believed that fourth or fifth order effects are really important, but in order to force the design to pay attention to intermediate levels of the number of children to avoid only comparing two extreme options. Fourth, although it seemed to be more realistic to include a multiplicative effect of the amount variable, this led to a lot of profiles with 1 or 12 children due to the D-optimality criterion. Therefore, an additive effect of the amount variable was used. Fifth, no alternative specific constants were needed to capture the mean effect of the unobserved factors in the error terms for each of the alternatives seen that the alternatives were unlabelled and that there was no opt-out option (Hensher et al., 2005b).

The model in Eq. 19 involves 153 parameters. This implies that many choice situations are needed which increases the cognitive burden for the respondents. A solution was to use an incomplete block design. This means that different sets of choice cards were created which were randomly distributed among the respondents in order to increase the number of choice situations. It was chosen to opt for 5 blocks of 12 profiles each or 120 options in total because each choice card consists of two mutually exclusive alternatives. The respondents were randomly assigned to one of the five blocks.

After examining the final choice sets, three adjustments needed to be added to exclude unrealistic choice cards. First, it was found that there were multiple profiles with extreme options (1 child and 12 children) in the design since SAS 9.4 can only create D-optimal designs. A solution was to force the SAS software to include profiles with at least three proportions that differ from zero. Second, there were a lot of choice sets where none of the children followed primary education. This was solved by excluding all profiles in which at least 50% of the children did not receive a primary education certificate seen that the drop-out ratio in primary schooling is 46% (The World Bank, 2020). Third, alternatives with an extremely high number of boys compared to girls (or vice versa) were excluded by adding the constraint that the probability of boys and girls under the binomial distribution had to be larger than 0.1 (Eq. 20).

Eq. 20
$$\binom{n}{k} 0.5^k * (1-0.5)^{n-k}$$

The final experimental design based on Equation 19 with the constraints stated above can be found in Table 8-3 in Appendix. This was obtained via the OPTEX procedure in the SAS 9.4 software which searches for optimal experimental designs by using a set of candidate design points and a specified model (SAS Institue Inc., 2014).

The choice cards themselves were created using Microsoft Powerpoint (Microsoft, 2010). Seen that the adult literature ratio is 39% (Central Statistical Agency & ICF International, 2016), visual aids were used (even though Couper, Tourangeau and Kenyon (2004) mention that these could affect the respondents' answers). The choice of the appropriate images was also investigated during the focus group discussions. A possible choice card is shown in Figure 3-8.

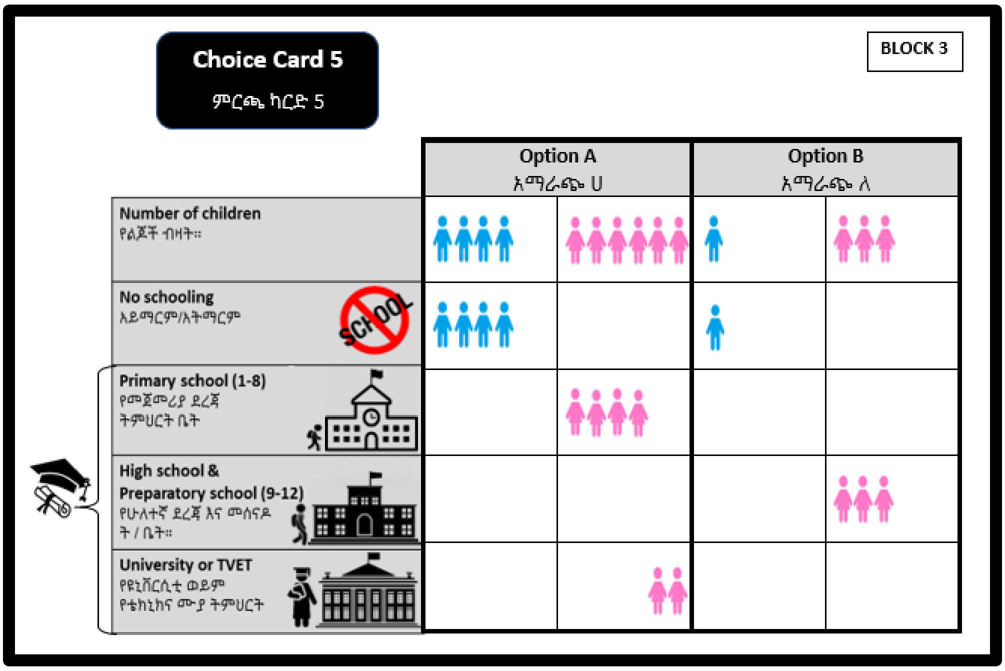


Figure 3-8: Example of one of the choice cards

3.3.3. Implementation

After the experiment was designed, it was implemented in the field. The implementation of the DCE is visualized in Figure 3-9. It consists of the sampling strategy, selection of the target group and the enumerators. Next, the implementation of the actual DCE and the data processing is described.

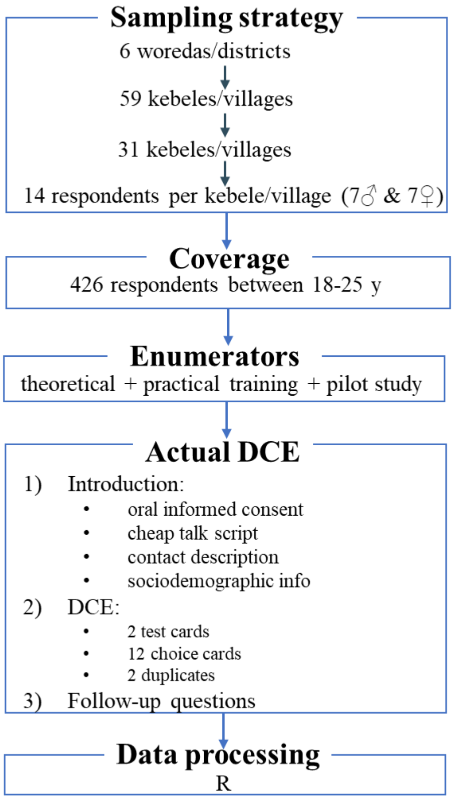


Figure 3-9:Visualization of the implementation of the DCE which includes the sampling strategy, followed by the selection of the target group, the enumerators, the performance of the actual DCE and the data processing.

For the sampling strategy, a three-stage sampling strategy was performed. This type of sampling strategy was used to reduce transportation costs and to increase fieldwork efficiency (de Leeuw et al., 2008; Roose & Meuleman, 2014). The sampling strategy started with the use of the sampling list from the household survey in Southern Ethiopia collected in the summer of 2018. In the first stage, 59 villages (i.e. kebele) from the 6 districts (i.e. woredas) were selected via stratification based on district and agroecology. For this research, in the second stage, 31 villages were randomly selected from the list of 59 villages of the 2018 household survey. In the third stage, a fixed number of 14 respondents per village were selected via a simple random sample with random replacements if necessary. This resulted in 434 targeted respondents. However, the total number of completed interviews was 426 of which 217 male respondents and 209 female respondents for which the reasons are listed in Table 3-3.

Table 3-3: Number of collected surveys with specification of the number of non-able sample units, overcoverage (i.e. the noneligible elements among the contacts) and the number of incomplete surveys

	Total	**Female**	**Male**
Total number of surveys	434	212	222
Number of non-able sample units	2	0	2
Cases of overcoverage	4	2	2
Number of incomplete surveys	2	1	1
Total number of respondents	426	209	217

The focus of this research is on ex ante fertility preferences (i.e. before respondents have completed their reproductive lifetime). This was chosen for two reasons. The first reason was to minimize bias due to rationalization by comparing the number with the current number of children (Bongaarts, 2011). The second reason was that fertility preferences are shown to vary significantly over individual's lifetime due to high existential uncertainty in SSA (Trinitapoli & Yeatman, 2018). Consequently, the coverage of the DCE included only men and women in the age category of 18-25 in the study area. The minimum age of 18 was chosen because it is forbidden to marry before this age, although the prevalence rate of child marriage is quite high (40% of women ages 20–24 years are married before the age of 18) (United Nations, 2019a). The maximum age was chosen at 25 based on the median age at first birth of 19.5 years and the median age at first marriage of 18 years (Ethiopian Society of Population Studies, 2008).

The DCE was performed in the period August-October 2019. Six interpreters were employed to collect the data. The recruitment was done using the requirements that the interpreters had to speak Amharic and at least some other languages (Gamo, Konso, Dirasha, Zayse, Gidicho, Masholle, Wolaytta or Oromo), and had some experience with data collection. To minimize processing errors, tablets were used to collect the data. This was done via the free software application 'Open Data Kit' (ODK Development Team, 2019). Before implementing the survey, the enumerators were trained in two ways: during a theoretical training of two days and a one-day practical training. First, the theoretical training sessions comprised an explication regarding interviewer techniques, a theoretical and practical explanation of the different parts of the survey (the non-response form, the DCE and the follow-up questions), a discussion on the concepts used in the DCE and a practical exercise with the tablets. Second, the practical training involved a one-day field experience during the pilot study. The objective of the training was to minimize the interviewer effects on the DCE implementation (de Leeuw et al., 2008). The pre-test or pilot study was performed with the interpreters in Shelle Mela, a village that was not part of the sampling list. The pilot study was done to ensure that the enumerators and the respondents understood the DCE well. Besides, the goal was to control if the attributes and the design of the choice cards were well chosen (in terms of question wording and visualization) to reduce measurement errors and to determine the time frame needed to perform all the DCE's.

Afterwards, there was a debriefing on their performance and modifications on the questionnaire were made based on the collected data. The participants were evaluated through their performance during the theoretical and practical sessions. Ultimately six enumerators were selected.

For the actual DCE, the respondents were asked to come to the agricultural training centre of the village where they were randomly assigned to an enumerator. Each enumerator was expected to perform around four to seven DCE's per day. Field supervision was undertaken to assure that the DCE was conducted correctly and to identify problems early in the data collection process and solve them as fast as possible. During the duration of the DCE, the quality of the data was checked, and specific feedback was given to the enumerators to improve their performance.

The DCE started with an oral informed consent (analogous to Figure 8-2 in Appendix). Thereafter, there was an introduction with a cheap talk script in which the objective, the method and the concepts of the used attributes were explained. Confidentiality and anonymity were assured. Additionally, consent was asked in order to reduce non-response errors (i.e. incapacity to obtain information for all sampled units on all items) because the questions that were asked could be rather sensitive (de Leeuw et al., 2008). The respondents were told that they could discontinue at any moment during the experiment they want. Measurement errors and hypothetical bias (i.e. inconsistent behaviour because of non-real choices that the respondents must make) were mitigated by including the short introduction before performing the DCE and by screening the data afterwards for implausible responses (Loomis, 2014). Subsequently, the choice experiment was performed. First, a trial of two choice cards including one dominant choice card was used to ensure that the respondents understood the task to eliminate the measurement errors. Next, the DCE was implemented with 12 choice cards. The order of the choice cards was randomized between different respondents in order to reduce measurement errors and the informational bias.

Afterwards, some follow-up questions were asked to the respondents. These include questions about their current marital status, their current and ideal number of children and knowledge about contraception because these are important determinants of fertility. Input was given by the thesis supervisors, the participants of the focus group discussions and the enumerators. An example of the answer sheet for these questions can be found in Figure 8-3 and 8-4 in Appendix. This was digitized in the tablets.

To minimize the non-response error, several steps were taken. First, to reduce **unit non response** (i.e. no response due to non-contact or refusal), a training[26] to customize the request to participate was used, next to a non-contact description form. The non-contact description included a detailed listing of the cases of overcoverage, number of non-able sample units and the number of incomplete surveys (Bethlehem, 2009). Secondly, to minimize **item non-response** (i.e. failure to obtain an answer on a specific choice card or question), the alternatives on the choice card were chosen to be as realistic as possible and confidence and anonymity was assured at the beginning of the DCE (de Leeuw et al., 2008).

Finally, the data processing included identifying inconsistencies and decoding of non-numeric questions. Data editing was accomplished using the R Software since the records of processing in this scripted language can be revised and rerun afterwards if necessary (R Core Team, 2020).

3.3.4. Analysis

The analysis of the DCE data started with the selection of an appropriate utility function that fits the data best (Figure 3-10). Therefore, 27 mixture-amount models were tested by putting them into the 3 discrete choice models that were used to reveal the preferences. Based on information criteria and the assumptions above, the best utility function was selected to analyse the data in order to answer the research questions.

Different types of discrete choice models can be used to reveal which attributes contribute to a chosen alternative. The following three models were used: the conditional logit model (CLM which assumes full preference homogeneity among respondents), the mixed logit model (MXLM which assumes full preference heterogeneity among respondents) and the latent class model (LCM which lies in between the CLM and MXLM assuming preference homogeneity among segments of the respondents). These will be explained in the sections below.

Then, before answering the research questions, a robustness check was performed. Next, the analysis of the heterogeneity in fertility preferences was done via the MXLM and the LCM. Afterwards, the optimal values of the attributes were calculated from an optimization perspective and the quantity-quality trade-off was analysed via the marginal rate of substitution.

[26] This is part the Leverage-Saliency theory (Groves et al., 2000). According to this theory, customizing the request to participate can increase the participation rate. Some design characteristics (e.g. topic, time…) can be judged by the interviewers as having a large leverage for the respondents. Then, the interviewers can tailor the introduction via focusing on the communication of features with a positive leverage, i.e. making them more salient.

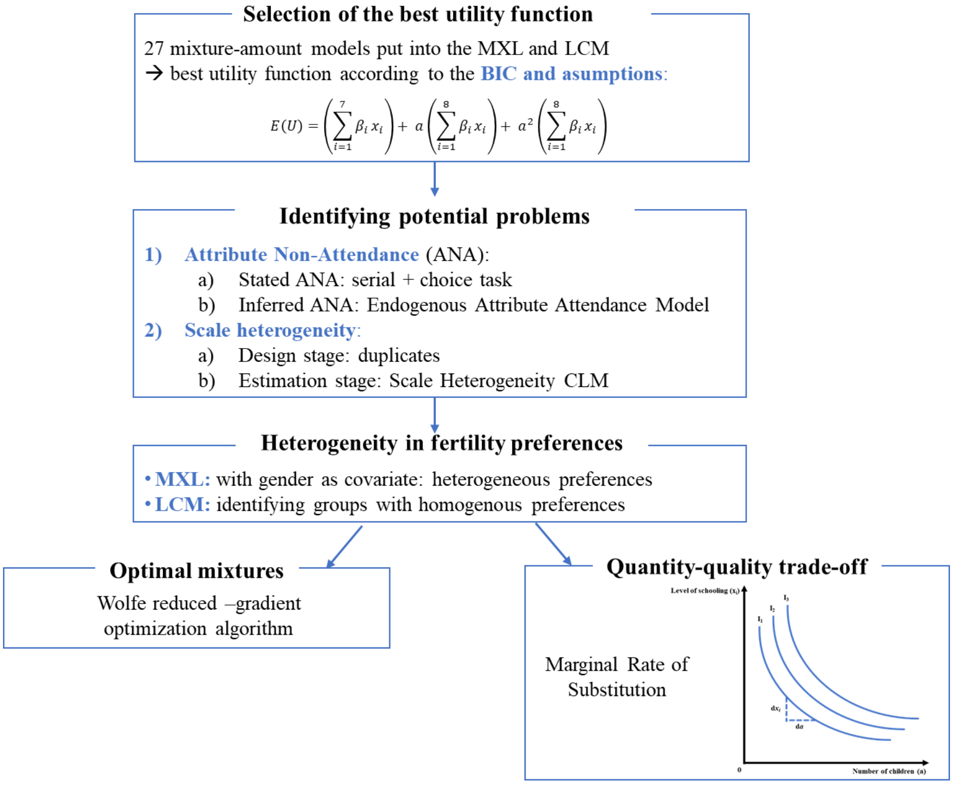


Figure 3-10: Visualization of the data analysis to answer the research objectives. This starts with the selection of the utility function that best fits the data. Then, the heterogeneity in fertility preferences is analyzed via the mixed logit model (MXLM) and the latent class model (LCM) followed by a subsequent analysis to account for potential problems. In addition, the quantity-quality trade-off was analysed via the marginal rate of substitution.

3.3.4.1. Theoretical framework of discrete choice models

<u>Conditional and multinomial logit model (CLM and MNLM)</u>

The CLM was first introduced by McFadden (1974). It is a common model for analysing consumer's demand of heterogeneous products. Recall Eq. 13 which becomes (Eq. 21):

Eq. 21 $$U_{ni} = Z_{ni}\boldsymbol{\beta} + \varepsilon_{ni}$$

where U_{ni} is the latent utility of individual *n* for alternative *i* in a certain choice set, Z_{ni} is a vector that differs across alternatives and possibly across individuals as well. β is a vector of parameters and ε is assumed to be independent and identically distributed (I.I.D.) following a Gumbel (Extreme Value Type 1) distribution. The CLM can be used to estimate the utilities as follows. Assume that individual *n* maximizes its utility over all the *i* alternatives, called *I*, available. Let y_{ni} be the choice of individual *n* over all the *i* alternatives, then (Eq. 22):

Eq. 22 $$y_n = \text{argmax}_i(U_{n0}, U_{n1}, U_{n2}, \dots, U_{nI})$$

Let δ_{ni} be the mean utility that person *n* receives from choosing alternative *I* (Eq. 23):

Eq. 23 $$\delta_{ni} = Z_{ni}\boldsymbol{\beta} = V_{ni}$$

Then, by integration over the Gumbel distribution, the probability that individual *n* chooses alternative *i* can be written in closed form as (Eq. 24):

Eq. 24 $$P(y_n = i) = P^n(i) = \frac{e^{\delta_{ni}}}{\sum_{j=1}^{I} e^{\delta_{nj}}} = \frac{e^{Z_{ni}\boldsymbol{\beta}}}{\sum_{j=1}^{I} e^{Z_{nj}\boldsymbol{\beta}}}$$

The derivation of this formula can be found in Cameron and Trivedi (2005).

Suppose now, instead of having variations in the vector Z_n, that this vector is the same for all alternatives, but has different effects across individuals. This model is called the multinomial logit model (MNLM) for which Eq. 1 can be written as (Eq. 25):

Eq. 25 $$U_{ni} = Z_n \boldsymbol{\beta_i} + \varepsilon_{ni}$$

where U_{nj} is the latent utility of individual *n* for alternative *i* ranging from 1 to I. β is a vector of parameters and ε is again assumed to be independent and identically distributed (I.I.D.) following a Gumbel (Extreme Value Type 1) distribution. So, MNLM and CLM are similar where MNLM is used if the model is defined in terms of the covariates whereas CLM is used if the model is modelled in terms of the attributes. In DCEs, the objective is to relate the choice to the attribute levels and therefore, it is more appropriate to use the term CLM, although both CLM and MNLM are used interchangeably in the choice experiment literature.

Nonetheless, this model has some disadvantages. First, adding another alternative or changing the characteristics of a third alternative (e.g. the attribute level for a choice experiment) does not affect the relative odds between 2 alternatives (Eq. 26):

Eq. 26 $$\frac{P^n(h)}{P^n(g)} = \frac{e^{\delta_{nh}}}{e^{\delta_{ng}}} = \frac{e^{Z_{nh}\boldsymbol{\beta}}}{e^{Z_{ng}\boldsymbol{\beta}}}$$

This property is called the independence from irrelevant alternatives (IIA). Hence, the estimated parameters will be biased and the substitution patterns are not very realistic (see for example the famous 'blue bus – red bus' example from McFadden (1974)[27]). Many other models have been proposed to alleviate the problem of IIA, such as generalized extreme value, multinomial probit and mixed logit models (Cameron & Trivedi, 2005; Revelt & Train, 1998). However, they often have other assumptions which can make the calculation computationally infeasible. A second disadvantage is that the number of choice sets have to be at least as large as the number of parameters in the model. A third disadvantage is the inability of the model to include taste heterogeneity between respondents with identical covariates. Due to the fact that this research wanted to reveal the heterogeneity of the respondent's choices, this model was not used in this research. Nevertheless, this model is relatively easy to calculate and has proven its applicability in literature.

[27] This example assumes that individuals have two modes of transportation: a red bus or a car. Suppose that a blue bus is added to the transportation options. Due to the property of IIA, the CLM predicts that the relative choice probability between the red bus and the car remains unchanged. However, this is not realistic, since the two buses are closer substitutes so people will probably substitute more from the red bus than from the car to the blue bus.

Mixed logit model (MXLM)

The MXLM combines the CLM (with alternative-variant regressors) and the MNLM (with alternative-invariant regressors) in order to eliminate the effect of IIA (Revelt & Train, 1998). So, the CLM/MNLM and MXLM can be seen as two ends of a continuum ranging from homogeneous to heterogeneous preferences. Again, recall Eq. 13 which becomes now (Eq. 27):

Eq. 27 $$U_{ni} = Z_{ni}\boldsymbol{\beta_n} + \varepsilon_{ni}$$

which assumes that respondents have heterogeneous preferences. This is reflected by the fact that every respondent has its own parameter vector β_n which is no longer fixed but follows a continuous distribution. Therefore, the parameter β_n can be seen as a pre-defined density function with the population level parameter estimates ϕ consisting of a mean β and a standard deviation γ_n (Eq. 28):

Eq. 28 $$U_{ni} = Z_{ni}(\boldsymbol{\beta} + \boldsymbol{\gamma_n}) + \varepsilon_{ni}$$

Here, the vector of coefficients β_n is assumed to be random and (log)normal distributed as $\pi(\beta|\phi)$ since both negative and positive preferences can be expected (Goos & Hamidouche, 2019). So, the MXLM assumes at the same time heterogeneous preferences among respondents, but these preferences are drawn from a population distribution $\pi(\beta|\phi)$ which limits the degree to which these preferences can differ between respondents. This results in the advantage that the number of choice sets for the respondents are irrespective of the number of parameters in the model. When applying the CLM or MNLM there was no such a restriction. To obtain the unconditional choice probability, a weighting factor is needed which is the weighted average of the following conditional probabilities via integration (Eq. 29):

Eq. 29 $$P(y_n = i) = P^n(i) = \frac{e^{\delta_{ni}}}{\sum_{j=1}^{I} e^{\delta_{nj}}} = \frac{e^{Z_{ni}\boldsymbol{\beta_n}}}{\sum_{j=1}^{I} e^{Z_{nj}\boldsymbol{\beta_n}}}$$

This results in the following unconditional probability of choosing one alternative (Eq. 30):

Eq. 30 $$P(y_n = i) = P^n(i) = \int \frac{e^{Z_{ni}\boldsymbol{\beta_n}}}{\sum_{j=1}^{I} e^{Z_{nj}\boldsymbol{\beta_n}}}\, \pi(\boldsymbol{\beta}|\boldsymbol{\phi})d\boldsymbol{\beta}$$

Latent class model (LCM)

The LCM can be used to capture the heterogeneity across segments of the population and not from individual respondents (Hall, 1999; Van Loo et al., 2018). This model assumes that there are a limited number of different segments in the group of respondents with a similar preference within one group. So, this model can be seen as lying in between the continuum of the MNLM and the MXLM. The LCM is a special case of the MXLM whereby the underlying continuous distribution of the vector of coefficients β_n becomes a discrete distribution. Although it makes the model less flexible, the sources of the preference heterogeneity can be identified and no assumptions about the distribution are needed anymore (Greene & Hensher, 2003). It is a type of unsupervised classification method which is model-based and therefore has advantages over other types of cluster analysis (Schreiber & Pekarik, 2014).

Every segment possesses its own parameter vector β_c. Eq. 13 takes now the following form (Eq. 31):

Eq. 31 $$\mathrm{U_{ni|c}} = \mathrm{Z_{ni}}\boldsymbol{\beta_c} + \varepsilon_{\mathrm{ni}}$$

So, within each class, the probability that individual *n* selects alternative *i* follows a CLM. Then, the unconditional choice probability that individual *n* selects alternative *i* belonging to the choice set ranging from 1 to I alternatives can be written as (Eq. 32):

Eq. 32 $$\mathrm{P(y_n = i) = P^n(i)} = \sum_{c=1}^{C} P_{ni|c}\, P_{nc}$$

where $P_{ni|c}$ is the probability that individual *n* from class *c* selects alternative *i* obtained via integration (Eq. 33):

Eq. 33 $$P_{ni|c} = \mathrm{P(y_n = i|c) = P^n(i|c)} = \frac{\mathrm{e}^{\delta_{\mathrm{ni}|c}}}{\sum_{\mathrm{j=1}}^{\mathrm{I}} \mathrm{e}^{\delta_{\mathrm{nj}|c}}} = \frac{\mathrm{e}^{\mathrm{Z_{ni}}\boldsymbol{\beta_c}}}{\sum_{\mathrm{j=1}}^{\mathrm{I}} \mathrm{e}^{\mathrm{Z_{nj}}\boldsymbol{\beta_c}}}$$

and P_{nc} is the probability that individual *n* belongs to class *c* (Eq. 34):

Eq. 34 $$P_{nc} = \mathrm{P(n \in c) = P^n(c)} = \frac{\mathrm{e}^{\theta_{\mathrm{c}} z_n}}{\sum_{\mathrm{c=1}}^{C} \mathrm{e}^{\theta_{\mathrm{c}} z_n}}$$

θ_c is a vector of parameters related to class-membership, z_n is an optional set of characteristics for individual *n* and *C* is the number of classes of which it is recommended to select this via the Akaike information criterion corrected for small sample sizes (AICc) or Bayesian information criterion (BIC) which are based on the log-likelihood (LL) (Greene & Hensher, 2003; Nylund-Gibson, 2007).

Choice models for mixtures

To model the choice experiment stated above, the systematic parts V_{ni} of the econometric models proposed above need to be replaced by the mixture models from section 3.3.2.2. However, due to property of the mixture constraint for mixture models (Eq. 13), an adjustment has to be made (Goos & Hamidouche, 2019). To see this, consider the following example. If the systematic part of the MXLM is replaced by linear terms including the proportions, such as for example the first order Scheffé model, this becomes (Eq. 35):

Eq. 35 $$E(U) = V_{\mathrm{ni}} = \mathrm{Z_{ni}}\boldsymbol{\beta_n} = \sum_{j=1}^{q} \beta_j x_{ji} = \sum_{j=1}^{q-1} \beta_j x_{ji} + \beta_q\left(1 - x_{1i} - \cdots - x_{(q-1)i}\right) = \beta_q + \sum_{j=1}^{q-1} (\beta_j - \beta_q) x_{ji}$$

with *x* the proportion of ingredient *j* in mixture *i*. Due to the mixture constraint, this expression contains a constant term, β_q. However, as described in section 3.3.1., relative differences in the systematic part of the utilities are important rather than absolute differences. Therefore, any choice model involving a constant term will be inestimable.

One solution to cope with this is to drop one of the linear terms in the ingredients (Goos & Hamidouche, 2019). This ingredient term can be interpreted as the default option with value 0. Then, all the other terms must be interpreted as relative to the reference ingredient. So, this solution will not affect the qualitative conclusions, only the quantitative conclusions. This technique is equal to the technique of dealing with dummy variable coding for qualitative factors in regression analysis where the number of dummy variables is always equal to the number of categorical variables minus one. For the example above involving the first order Scheffé model, this becomes (Eq. 36):

Eq. 36

$$E(U) = V_{\mathrm{ni}} = \mathrm{Z}_{\mathrm{ni}}\boldsymbol{\beta} = \sum_{j=1}^{q-1}(\beta_j - \beta_q)x_{ji}$$

The same applies for higher order Scheffé models and the other choice models (MNLM, MXLM and LCM). For example, for the final Scheffé model from Eq. 19 used for the design of the choice experiment, this results in the following expression for a MXLM choice model (Eq. 37):

Eq. 37

$$E(U) = V_{\mathrm{ni}} = \mathrm{Z}_{\mathrm{ni}}\boldsymbol{\beta}_{\mathbf{n}} = \sum_{k=1}^{7}(\beta_k - \beta_8)x_i + \sum_{k=1}^{7}\sum_{l=k+1}^{8}\beta_{kl}x_{ki}x_{li} + \sum_{k=1}^{6}\sum_{l=k+1}^{7}\sum_{m=l+1}^{8}\beta_{klm}x_{ki}x_{li}x_{mi} + \gamma a + \gamma^2 a + +\gamma^3 a + \gamma^4 a + \gamma^5 a$$

3.3.4.2. Selection of the best utility function

To select the most suitable utility function for the collected data, it is important to test various models (Van Der Pol et al., 2014). Since the experimental design contained 5 blocks of 12 choice sets with each 2 alternatives, a maximum of 120 parameters can be estimated. Therefore, only 27 different possible Scheffé models could be tested. These are indicated in grey in Table 3-4. To select a model that best fits the data, different quality measures were used to compare the models from Table 3-4. The most used information criteria are the AICc (Sugiura, 1978) and the BIC (Schwarz, 1978). AICc and BIC are in-sample criteria based on the log-likelihood (LL). Both criteria make a trade-off between goodness-of-fit and parsimony where BIC penalizes more for more complex models which is recommended by Schwarz (1978). The BIC decreases with an increase in the number of parameters (Figure 3-11). A 10-point decrease indicates a highly significant improvement (Kass & Raftery, 1995). For the additive effect of the amount variable, a maximum was put at power eight since the BIC did not improve further. Next, other more recent models than the Scheffé models as stated above (4.3.2.2.) were also tested. However, these did not lead to an improvement of the best Scheffé models. Therefore, these are not shown here.

Table 3-4: Number of parameters of the different discrete choice models

Amount effect		First order or linear	Second order or quadratic	Third order or special cubic	Fourth order or full cubic
None		7	35	147	175
Additive	*Linear*	8	36	148	176
	Quadratic	9	37	149	177
	3rd power	10	38	150	178
	4th power	11	39	151	179
	5th power	12	40	152	180
	6th power	13	41	153	181
	7th power	14	42	154	182
	8th power	15	43	155	183
Multiplicative	*Linear*	15	70	294	350
	Quadratic	23	105	441	525
	3rd power	31	140	588	700
	4th power	39	175	735	875
	5th power	47	210	882	1050
	6th power	55	245	1029	1225
	7th power	63	280	1176	1400

Besides, the model selection needs to be based on good interpretability of the model (Swait, 1994) and theoretical prior information (Ruto et al., 2007) in order not to add superfluous parameters to the model. Concerning theoretical prior information based on literature, two assumptions were made:

a) The relationship between the utility and the number of children first increases until a maximum and then stagnates or decreases, based on the normal distribution of family sizes among the world (The World Bank, 2020). So, the model has to include a quadratic or higher order effect for the variable *number of children*.
b) Interaction effects between the number of children and the schooling components seem to be more realistic (Burks et al., 2019; Mogstad & Wiswall, 2009).

Therefore, a multiplicative model with the variable *number of children* of at least power two was chosen. From the remaining possible models, the one with the best BIC criterion was selected. However, for a few number of classes of the LCM, no quality measures could be calculated due to a non-symmetric or highly singular variance matrix. Table 8-4 in Appendix provides the information criteria for all the possible Scheffé models.

For the MXLM, the first order Scheffé model has lower BIC values compared with the second order Scheffé model. For the first order models, Figure 3-11 shows a significant improvement in terms of the BIC for a quadratic effect of the *number of children*. Also, for the LCMs, this model performed the best for five classes. For the LCM with two classes, a first order multiplicative model of order five performed best. For the LCM with three classes, a first order multiplicative model of order three resulted in the lowest BIC. Nonetheless, according to the Pareto principle, a simpler model was preferred (Goos & Jones, 2011). For the LCM with four and six classes, a first order multiplicative model of order one was the best but this does not fulfil our first assumption. Therefore, the first order multiplicative model with a quadratic effect of the amount variable was chosen. The form of this model is as follows (Eq. 38):

Eq. 38
$$E(U) = \left(\sum_{i=1}^{7} \beta_i \, x_i\right) + \, a\left(\sum_{i=1}^{8} \beta_i \, x_i\right) + \, a^2\left(\sum_{i=1}^{8} \beta_i \, x_i\right)$$

This corresponds with the LCM of five classes. For the MXLM this model involves 23 preference parameters. For the LCM, this involves 119 parameters (23 times the number of classes plus the number of classes minus one (because the class probabilities need to be estimated as well)) (Goos & Hamidouche, 2019).

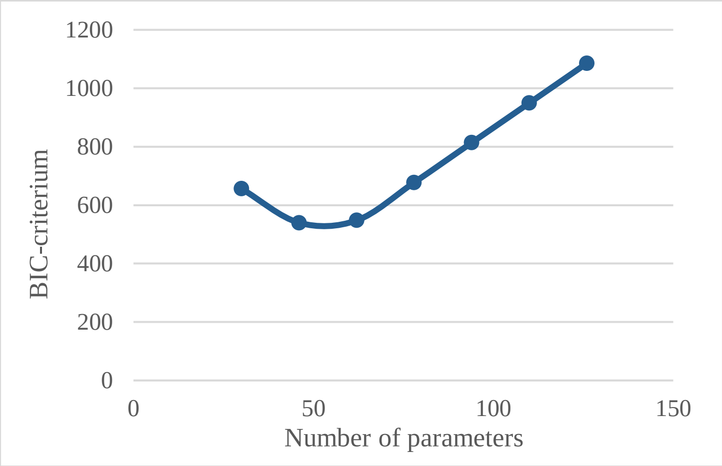


Figure 3-11: Effect of the number of parameters on the BIC-criterion for the first order or linear Scheffé model with a quadrative amount-effect

3.3.4.3. Robustness checks

Several problems can be encountered when performing a DCE of which the most serious are attribute non-attendance (ANA) and scale heterogeneity.

<u>Attribute non-attendance</u>

Attribute non-attendance occurs when the respondents do not consider a certain attribute in making a trade-off between the alternatives (Hensher, 2008). Three solutions are proposed to take this into account. The **first solution** is asking the respondents after the choice experiment (i.e. serial stated ANA) or after each choice card (i.e. choice task stated ANA) to indicate which attribute they did not consider in making the trade-off (Campbell et al., 2008; Hensher et al., 2005a; Scarpa et al., 2010). These results can be included in the model via a conventional or validation ANA model (Caputo et al., 2018). A **second solution** is inferring ANA during the data analysis (i.e. inferred ANA). This can be done via four methods. A first approach is an equality constrained LCM (Hensher & Greene, 2010; Scarpa et al., 2009; Van Loo et al., 2018). Secondly, this can be done via the combined latent class mixed logit (Hess et al., 2012) or random parameter mixed panel logit models (Hess & Hensher, 2010). Thirdly, Kehlbacher, Balcome and Bennett (2013) propose a shrinking approach. Fourth, Hole (2011) put forward an endogenous attribute attendance approach. A **third solution** is revealing ANA via eye tracking (Van Loo et al., 2018).

In this survey, three methods were used. First, serial stated ANA was performed. Based on these results, an MXLM model with dummy variables of the most ignored attributes was compared with the 'standard' MXLM since the 'standard' MXLM also accounts for preference heterogeneity. Second, choice task stated attribute attendance was queried. Respondents were asked to think aloud while making their decisions in order to reduce the risk of respondent fatigue which can result in more random decision-making (Campbell et al., 2015; Ryan et al., 2009). Third, ANA was taken into account using the endogenous attribute non-attendance logit model (EAAlogit) via the 'eaalogit' command in Stata (Hole, 2010). This model was chosen since it is able to include all possible attribute subsets as opposed to a limited number. Since the EAAlogit could not estimate the full model with the main effects, the non-attendance pattern was gradually build up which is the reason that only three gamma coefficients are shown in the output. Further, the EAAlogit model assumes preference homogeneity. Therefore this was compared with the 'standard' CLM which also assumes preference homogeneity as stated above. The CLM can be estimated using the common maximum likelihood approach whereby the estimated parameter values are the ones that maximize the likelihood of observing the data (Cameron & Trivedi, 2005). However, if the data shows separation, the maximum likelihood estimation does not exist or is biased when the sample is small. Therefore, the Firth penalized-likelihood estimation procedure was used which overcomes these problems and is recommended by Kessels, Jones & Goos (2019).

Scale heterogeneity

In the CLM, MXLM and LCM described above, the error variance is assumed to be constant across respondents. In reality, it is possible that this differs across respondents due to a different ability to understand and perform the DCE and due to a different commitment to the DCE (Hess & Stathopoulos, 2013; Hossein et al., 2018; Louviere et al., 2002). This is called scale heterogeneity. A solution to attenuate scale heterogeneity can occur in two stages of the analysis. It can occur in the **design stage** by choosing the appropriate number of alternatives and attributes, by including duplicates of the choice cards or by different experimental designs or formatting for different groups of respondents. A second option occurs in the **estimation stage** by including scale heterogeneity in the model via a parametrization of the scale factor (Hess & Train, 2017), controlling for it using a heteroscedastic logit model (e.g. scaled heterogeneity CLM, MXLM or

LCM or generalized CLM, MXLM or LCM) (Collins et al., 2012; Deshazo & Fermo, 2002; Fiebig et al., 2010), avoiding the heterogeneity by using individual-level discrete choice models (Louviere et al., 2008; Melstrom et al., 2017), via a coefficient plot (Wright et al., 2018), the Swait and Louviere Test (Wright et al., 2018) or a scale adjusted LCM (Wright et al., 2018).

To account for scale heterogeneity in this DCE, during the **design stage** it was chosen to put only two alternatives per choice card to reduce the cognitive burden for the respondents. Moreover, two duplicates of the twelve choice cards were added after the DCE itself. Next, to infer for scale heterogeneity during the **estimation stage**, this was analysed via the scaled MXLM (also called the generalized multinomial logit model II) with the 'gmnl' command in Stata (Gu et al., 2013). This was then compared with the 'standard' MXLM to analyse the difference between a model that only accounts for preference heterogeneity (i.e. the 'standard' MXLM) and a model that accounts for both preference and scale heterogeneity (i.e. the scaled MXLM).

3.3.4.4. Analysis of the heterogeneity in fertility preferences

To analyse the (gendered) heterogeneity in fertility preferences via the DCE, two models were used: the MXLM and the LCM. To take into account the possible discrepancy in preferences between male and female respondents, as recently argued by Ndagurwa (2019), the interaction term with the covariate 'gender' was added to the MXLM. Furthermore, the LCM was implemented to look for groups of respondents characterized by socio-demographic characteristics beyond gender. Then, these characteristics could be used to make more specific policy recommendations for groups of respondents with a similar preference.

For the MXLM, in the literature, two types of estimation procedures are most commonly used: maximum likelihood estimation and hierarchical Bayes estimation (J. Huber & Train, 2001). Using the maximum likelihood procedure resulted in convergence failure for higher-order models and large number of random draws were needed resulting in a larger computational burden and time (about 20 hours on average per model with a 4010U Processor) with the Stata Software. Therefore, it was opted to use the hierarchical Bayes estimation in the JMP Pro 14 software (SAS Institute Inc., 2018) which is asymptotically equivalent to the maximum likelihood estimator (J. Huber & Train, 2001). The hierarchical Bayes estimation procedure estimates the parameters of the posterior distribution via the Adaptive Bayesian method and Metropolis-Hastings approaches while accounting for uncertainty as explained by Train (2001) and Proust (2018). This was done for 10,000 iterations and 5,000 burn-in iterations. Non-Firth maximum likelihood estimates were used as starting values for the hierarchical Bayes algorithm. The posterior mean and the confidence limits are reported. The posterior mean is the average of the means of the subject-specific coefficient estimates whereas the confidence limits are constructed based on the 2.5 and 97.5 quantiles of the posterior distribution assuming a normal distribution.

LCMs are mostly estimated via the maximum likelihood estimation, (quasi-) Newton methods, the Berndt–Hall–Hall–Hausman method or the expectation-maximization algorithm. It was chosen to use the expectation-maximization algorithm via the ‘lclogit2’ command in Stata (Hong Il Yoo, 2019) since this was able to estimate non-linear models, allows for continuous responses, guarantees numerical stability and convergence to a global maximum (even for a large number of classes) and it has a faster run time and stability (Bhat, 1997; Train, 2008). Functionalities for LC analysis in the R, NLogit, Latent Gold and JMP software do not provide this algorithm. To ensure to reach a global optimum instead of a local maximum (Train, 2008), 50 different alternative starting values were tested. Standard errors and confidence intervals were obtained via the user-written ‘gllamm’ package (Rabe-Hesketh et al., 2002). No covariates were added to the model since this could introduce biases (Wurpts & Geiser, 2014). Instead, the sociodemographic differences between the classes were analysed afterwards in R (R Core Team, 2020).

3.3.4.5. Optimal mixtures

From an optimization perspective, it is interesting to calculate the respondents’ maximized utility for the attributes. This involved two steps since our model is a mixture-amount model. First, the maximized desirability of the number of children (*the amount*) was calculated for each of the estimated MXLM and LCM via the utility profiler in JMP Pro 14 (SAS Institute Inc., 2013). The maximized desirability of the utility is defined by Derringer & Suich (1980) (Eq. 39):

Eq. 39

$$d_i(\hat{Y}_i) = \begin{cases} 0 & if\ \hat{Y}_i < Y_{min} \\ \left(\frac{\hat{Y}_i - Y_{min}}{Y_{max} - Y_{min}}\right)^r & if\ Y_{min} < \hat{Y}_i < Y_{max} \\ 1 & if\ \hat{Y}_i > Y_{max} \end{cases}$$

where $d_i(\hat{Y}_i)$ is the individual desirability of response Y_i. The desirability function was calculated by interpolating cubics between the minimum and maximum utility values for the attribute number of children via the Wolfe reduced-gradient optimization algorithm. Second, an area plot was made for the eight schooling components (*the mixture*) in function of the number of children at the highest utility. The optimal mixture could then be found by looking at the corresponding schooling proportions at the maximized desirability value of the number of children.

3.3.4.6. The quantity-quality trade-off

To investigate the trade-off that respondents make between the quantity (i.e. number of children) and the quality (i.e. education) of child-raising, different techniques exist. Most studies compare the size of the parameter estimates of the attributes. However, the size can vary due to the relative impact of an attribute on the utility or due to the different utility scale of the attributes (Lancsar et al., 2007). Further, Lancsar et al. (2007) proposed six different techniques for DCEs: i) partial log-likelihood analysis, ii) the Hicksian welfare measure, iii) probability analysis, iv) best-worst attribute scaling, v) relative attribute importance and vi) the marginal rate of substitution (MRS). In this research it was chosen to use the MRS since this is the only measure which can handle non-orthogonal designs (in contrast to i), no predefined attribute changes or selection (as opposed to ii, iii and iv) and focus on the homogeneity in respondents' preferences (in contrast to v). The MRS calculates the willingness to reduce one unit of a certain continuous attribute for an increase of another continuous attribute and is scale independent. This is calculated as (Eq. 40):

Eq. 40 $$MRS_{a,x_i} = \frac{\partial U/\partial a}{\partial U/\partial x_i} = \frac{-dx_i}{da}$$

with U the utility, x_i a certain level of schooling (as defined in Table 3-2) and a the number of children. This was then visualized via indifference curves in JMP Pro 14 (SAS Institute Inc., 2013). Indifference curves are curves along which utilities are held constant implying that the respondent is indifferent between certain combinations of attributes (Mühlbacher et al., 2015). An example is provided in Figure 3-12. The curves' shape reflects individual tastes (defined by ethnicity, religion, age, etc.).

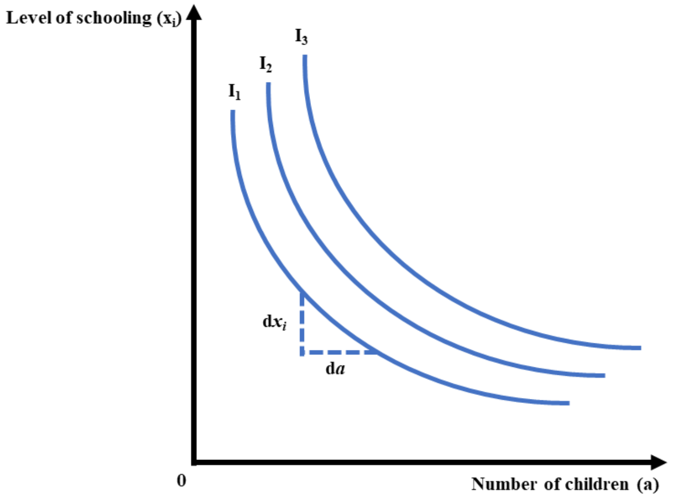


Figure 3-12: Indifference curves I1, I2 and I3 (in order of increasing fixed utility levels) for the number of children (a) and the level of schooling (x_i as defined in Table 3-2)

3.4. Follow-up questions

The analysis of the follow-up questions was performed with the R software (R Core Team, 2020). The sociodemographic characteristics were used to analyse the (gendered) heterogeneity in fertility preferences and to compare the characteristics from our sample with the national data. For ratio data, first the assumptions of parametric tests were analysed. Normality was checked via the standard Shapiro-Wilk test and homoscedasticity via the Levene's test in the package 'lawstat' (Gastwirth et al., 2019). If both assumptions were fulfilled, a two sample t-test was performed if differences between two groups were analysed or an ANOVA test with post-hoc Tukey test if the test involved differences between more than two groups. If the normality and/or homoscedasticity assumption of the parametric tests did not hold, a standard Mann-Whitney test was performed if there were two groups or a Kruskal-Wallis test with post-hoc Dunn's test (to adjust for the False Discovery Rate (Benjamini & Hochberg, 1995)) if there were more than two groups via the package 'FSA' (Ogle et al., 2020). For nominal data, differences between groups were analysed via the pairwise Chi Square test accounting for the False Discovery Rate via the package 'rcompanion' (Mangiafico, 2020).

4. Results

4.1. Robustness checks

4.1.1. Attribute non-attendance

Concerning **serial stated ANA**, Table 4-1 shows that 84% of the respondents reported that they considered all attributes during the choice experiment. In addition, 47% of the respondents stated that they found all attributes equally important. The attributes ‘gender of the children’ and ‘no schooling’ are reported to be somewhat ignored (by 7% and 7% of the respondents respectively) and least important (20% vs 27% respectively). These two possibly ignored attributes were therefore also analysed via a MXLM with the addition of dummy variables for these two attributes (Table 4-3). This reveals that, compared with the ‘standard’ MXLM, respondents who indicated that they did not consider certain attributes in making their choices have almost no significantly different preferences (indicated by the interaction terms of the dummy variables and the attributes). This holds except for respondents who stated that they had ignored the attribute gender, they have an even higher preference for a higher number of girls send to primary, secondary and tertiary education and boys send to tertiary education. Further, they have a lower preference for the number of children. Respondents who indicated to have ignored the attribute no schooling, have a higher preference for the number of girls send to tertiary education and boys to secondary school. Besides, they have a lower preference for the number of children. This reveals that, according to serial stated ANA, there is only some minor evidence for ANA.

Looking at the results of **choice task stated ANA** (Table 4-2), 40% of the options on the choice cards were made based on gender and educational level. 26% of the respondents explicitly took the educational level into account and 23% of the options was decided upon based on gender differences of the alternatives. This corresponds with the results of serial stated ANA since the attributes that are considered as the most important are also reported as the least ignored/least important, except for the number of children.

Inferred ANA via the EAAlogit model reveals that the significance levels, the sign and the magnitude of the variables are similar when compared with the ‘standard’ CLM (Table 4-3). However, three gamma coefficients (γ) are significant, indicating that there is some evidence for ANA. This is the case for the FTertiary, MTertiary and NChildren with corresponding probabilities of ANA of 19%, 21% and 35% respectively (based on the gamma coefficients; calculations not shown here). So, mainly the attributes FTertiary, MTertiary and NChildren are possibly affected by the ANA. Therefore, the coefficients of these attributes need to be interpret with caution. Nevertheless, for the EAAlogit model, the effects of FTertiary, MTertiary and

NChildren increase and the AICc and BIC improve compared to the 'standard' CLM. So, the possible ANA bias seems to be no problem.

Overall, there is some minor evidence for ANA. However, there is little concordance between the results of stated and inferred ANA. The reason could be that the design of the DCE is based on a mixture model which makes it very hard to ignore one of the attributes due to the collinearity (i.e. linear dependence) of the variables. For example, if it is known that a certain percentage of the children is send to a particular educational level, it can be inferred that the rest of the children is send to one of the other educational levels since the sum has to be equal to the total number of children. Therefore, it is assumed that our experiment is quite robust to ANA and it is not taken into account in the further analysis.

Table 4-1: Serial stated ANA after the choice experiment whereby respondents were asked which attribute they had ignored (1st column) and which they found the least important (2nd column)

Attributes	Ignored			Least important			
	Total	Female	Male	Total	Female	Male	
Sex of the children (%)	6.57	7.37	5.74	19.72	20.74	18.66	
Number of the children (%)	1.64	2.30	0.96	5.87	6.45	5.26	
No schooling (%)	6.81	6.45	7.18	27.00	22.12	32.06	**
Educational level (%)	0.23	0.00	0.48	0.23	0.46	0.00	
All attributes are considered/equally important (%)	83.98	84.33	85.65	47.18	50.23	44.02	

*Note: Comparison between the male and female subgroups was based the pairwise Chi Square test accounting for the False Discovery Rate. Significant differences of the mean of the male vs. female subgroups are shown with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

Table 4-2: Choice task stated attribute attendance (i.e. which attributes the respondents took into account for making each choice)

Attributes	Total	Female	Male
Sex of the children (%)	22.57	21.85	23.27
Number of children (%)	3.74	3.35	4.11
Educational level (%)	26.37	27.11	25.65
Sex of the children and number of children (%)	1.45	1.83	1.08
Sex of the children and educational level (%)	39.69	39.79	39.59
Number of children and educational level (%)	4.36	4.47	4.26
Sex of the children, number of children and educational level (%)	1.82	1.59	2.04

*Note: Comparison between the male and female subgroups was based the pairwise Chi Square test accounting for the False Discovery Rate. Significant differences of the mean of the male vs. female subgroups are shown with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

Table 4-3: Parameter estimates of the 'standard' MXLM obtained via the hierarchical Bayes estimator and the 'standard' CLM obtained via the Firth-Bias adjusted estimator for comparison with the models to account for ANA

Variable	MXLM				Serial stated ANA				Serial stated ANA				CLM				Inferred ANA via EAAlogit			
	Posterior Mean		Lower CI 95%	Upper CI 95%	Mean		Lower CI 95%	Upper CI 95%	Mean		Lower CI 95%	Upper CI 95%	Mean		Lower CI 95%	Upper CI 95%	Mean		Lower CI 95%	Upper CI 95%
FPrimary	1.38	*	0.90	1.93	1,20	*	0,63	1,73	1,51	*	0,94	2,02	0.91	*	0.53	1.29	0.94	*	0.55	1.33
FSecondary	2.40	*	1.88	2.90	2,30	*	1,75	2,77	2,55	*	1,98	3,10	1.54	*	1.14	1.94	1.80	*	1.37	2.23
FTertiary	4.14	*	3.56	4.70	3,90	*	3,31	4,47	4,26	*	3,70	4,81	2.68	*	2.29	3.08	3.82	*	3.25	4.39
MNoPrimary	1.45	*	0.87	2.00	1,44	*	0,88	2,07	1,60	*	0,98	2,17	0.86	*	0.43	1.29	1.12	*	0.63	1.61
MPrimary	0.89	*	0.40	1.40	0,72	*	0,20	1,23	1,02	*	0,43	1,59	0.46	*	0.04	0.87	0.46	*	0.03	0.90
MSecondary	1.78	*	1.30	2.21	1,65	*	1,15	2,14	1,83	*	1,21	2,36	1.35	*	0.96	1.75	1.30	*	0.90	1.71
MTertiary	5.56	*	4.84	6.30	5,16	*	4,47	5,80	5,88	*	5,22	6,59	3.77	*	3.32	4.24	5.63	*	5.02	6.23
NChildren	-0.24	*	-0.26	-0.22	-0,24	*	-0,26	-0,21	-0,24	*	-0,27	-0,22	-0.15	*	-0.16	-0.14	-0.29	*	-0.32	-0.26
FPrimary*Dummy_gender					27,17	*	11,56	42,79												
FSecondary*Dummy_gender					12,11	*	0,36	21,95												
FTertiary*Dummy_gender					59,55	*	43,68	74,50												
MNoPrimary*Dummy_gender					9,53		-14,36	32,61												
MPrimary*Dummy_gender					10,48		-5,81	25,00												
MSecondary*Dummy_gender					7,73		-14,18	22,91												
MTertiary*Dummy_gender					90,51	*	73,93	111,81												
NChildren*Dummy_gender					-3,26	*	-5,16	-1,83												
FPrimary*Dummy_no_schooling									-3,68		-10,42	3,77								
FSecondary*Dummy_no_schooling									5,14		-1,64	11,73								
FTertiary*Dummy_no_schooling									17,05	*	8,05	26,91								
MNoPrimary*Dummy_no_schooling									-1,67		-14,05	11,16								
MPrimary*Dummy_no_schooling									-6,81		-16,33	1,54								
MSecondary*Dummy_no_schooling									8,96	*	0,62	19,60								
MTertiary*Dummy_no_schooling									0,66		-9,29	9,92								
NChildren*Dummy_no_schooling									-1,17	*	-1,87	-0,57								
γ (FTertiary)																	1.44	*	0.72	2.16
γ (MTertiary)																	1.31	*	0.91	1.70
γ (NChildren)																	0.63	*	0.36	0.90
AICc			4230.75				4071,92				4056,04				5646.40				5425.15	
BIC			4283.05				4176,47				4160,60				5698.69				5497.06	
-2LL			4214.72				4039,81				4023,94				5592.19				5403.10	

*Note: Significant differences of the (posterior) mean on a 5% level are indicated with *. The first and second model that accounts for ANA includes interaction effects of the dummy variables gender and no schooling respectively (Ignored = 1; Not ignored = 0) with the main effects of the attributes. The third model takes ANA into account via the endogenous attribute attendance model (EAAlogit) where γ indicates the gamma coefficient used to calculate the probability of ANA. Since the EAAlogit model was not able to estimate the full model with the main effects, the non-attendance pattern was gradually build up. This is the reason that only three gamma coefficients are shown in the output. The explanation of the variables can be found in Table 3-2.*

4.1.1. Scale heterogeneity

First, scale heterogeneity was taken into account in the **design stage** by adding two duplicates of the choice cards to the DCE. The results indicate that 90% of the respondents gave the same answer on the duplicated choice card. Table 4-4 shows that, when compared to the 'standard' MXLM, the respondents who gave different answers on the duplicates have different preferences for FTertiary, MNoPrimary and MTertiary (indicated by the interaction terms of the dummy variables and the attributes). This suggests that there is some evidence for scale heterogeneity.

Second, scale heterogeneity was analysed during the **estimation stage**. The results of the scaled MXLM show that the significance, sign and magnitude of the variables are similar to the 'standard' MXLM (Table 4-4). However, the scale parameter (τ) is significant which provides evidence for some scale heterogeneity. Nevertheless, if the scale parameter is included in the model, the AICc and BIC do not improve for the scaled MXLM compared to the 'standard' MXLM.

Consequently, there is some minor indication for scale heterogeneity. However, this was accounted for in the design by starting the DCE with a trial of two test cards to minimize the differences in capability to understand the task. In addition, Hess and Rose (2012) argue that it is not completely possible to disentangle preference and scale heterogeneity with the current models. Therefore, it is not taken into account in the further analysis.

Table 4-4: Parameter estimates of the 'standard' MXLM obtained via the hierarchical Bayes estimator for comparison with the models to account for scale heterogeneity.

Variable	MXLM				MXLM with interaction of duplicates				Scaled MXLM			
	Posterior Mean		Lower CI 95%	Upper CI 95%	Mean		Lower CI 95%	Upper CI 95%	Mean		Lower CI 95%	Upper CI 95%
FPrimary	1.38	*	0.90	1.93	1.59	*	1.09	2.08	2.11	*	1.33	2.89
FSecondary	2.40	*	1.88	2.90	2.61	*	2.06	3.15	2.87	*	2.05	3.70
FTertiary	4.14	*	3.56	4.70	4.28	*	3.70	4.86	4.82	*	3.68	5.96
MNoPrimary	1.45	*	0.87	2.00	1.32	*	0.70	1.95	1.84	*	1.07	2.60
MPrimary	0.89	*	0.40	1.40	0.98	*	0.45	1.51	1.42	*	0.70	2.13
MSecondary	1.78	*	1.30	2.21	1.91	*	1.39	2.44	2.47	*	1.69	3.24
MTertiary	5.56	*	4.84	6.30	5.76	*	5.04	6.47	6.77	*	5.34	8.20
NChildren	-0.24	*	-0.26	-0.22	-0.25	*	-0.28	-0.23	-0.24	*	-0.29	-0.19
τ									0.83	*	0.62	1.03
FPrimary*Dummy_duplicate					-0.40		-2.39	1.60				
FSecondary*Dummy_duplicate					0.92		-0.87	2.71				
FTertiary*Dummy_duplicate					3.81	*	1.45	6.18				
MNoPrimary*Dummy_duplicate					4.42	*	1.56	7.28				
MPrimary*Dummy_duplicate					0.05		-2.17	2.26				
MSecondary*Dummy_duplicate					0.23		-1.94	2.41				
MTertiary*Dummy_duplicate					4.86	*	1.72	8.00				
NChildren*Dummy_duplicate					-0.18		-0.31	-0.06				
AICc			4230.75				3993.50				5433.77	
BIC			4283.05				4098.06				5492.61	
-2LL			4214.72				3961.40				-5415.74	

*Note: Significant differences of the mean of the posterior mean on a 5% level are indicated with *. The first model that accounts for scale heterogeneity includes interaction effects of the dummy variable for the answer on the duplicate (Different answer = 1; Same answer = 0) with the main effects of the attributes. The second model takes scale heterogeneity into account via the scaled MXLM where tau is the scale parameter. The explanation of the variables can be found in Table 3-2.*

4.2. Fertility preferences of male and female respondents (MXLM)

4.2.1. Sociodemographic characteristics

Respondents are on average 21 years old, 64% is Protestant and 63% have completed high school (Table 4-5). Respondents are mainly student (36%), have a farming-related job (22%) or a non-farming related job (21%). The large majority of respondents is not married (77%) and has no children (80%) (Table 4-6). Respondents with children, have on average two children. The ideal FSP is on average 4 children. During the focus group discussions, it was reported that there is a trend of declining fertility rates over the generations. The reasons for preferring a larger number of children by the previous generations were a sign of wealth, continuation of the ethnicity lineage and infant mortality. Moreover, respondents have on average a higher preference for boys than for girls (a preferred sex ratio of 1.118) which was also reported during the focus group discussions. Concerning the ideal level of education, on average respondents prefer a university degree for both their future sons and daughters. This was also stated during the focus group discussions. Schooling for boys was preferred over girls because investment in girls is perceived as lost since the success of the girl will go to the family of her future husband. Additionally, the focus group discussions revealed that Alternative Basic Education is not common in the study area. Furthermore, 88% of the respondents have knowledge about contraception (88%) of which he/she prefers and knows most of implants, depo and the male condom. About half of the respondents have not used contraception yet (51%) and 95% has access to contraception in the village. On average, 39% stated that there is never a shortage of contraceptives in the village whereas 30% indicated that there is sometimes a shortage. During the focus group discussions, it was also observed that contraception was well-known by most of the respondents, but the demand is quite low seen that the respondents are in their reproductive age or due to shortages.

Male and female respondents differ by age (on average 22 years vs 20 years respectively). Other discrepancies are that 59% of men vs 68% belong to Protestantism, 29% of men vs 34% of women have completed high school. In addition, 30% of men vs 13% of women are a family labourer on the own farm, 5% of men vs 18% of women are a family labourer in the household and 3% of men vs 0.5% of women a hired servant. (Table 4-5). The trend of significantly more women having high school as their highest educational level was also noted during the focus group discussions. The main reasons were economic and the belief in the community that girls will become married after the 8th or 10th grade. Additionally, more men (27%) are currently married than women (19%) (Table 4-6). During the focus group discussions, both men and women reported that child-raising was the main reason to become married because of a social sanction due to premarital childbearing because of religious reasons. Next, the age at first marriage in this survey is significantly higher for male respondents (20 years) than for female respondents (18 years). For respondents who are not married, the preferred age to become married is also significantly higher for men (27 years) than for women (24 years). A discrepancy between the age at first birth among respondents with children (20 years for men and 19 years for women) and the preferred age at first birth among respondents without children (28 years for men and 25 years for women) was also observed. Further, women have a significantly higher preference for university for their future sons (98% vs 94%). Besides, men stated more frequently to have knowledge about contraception than women (88% vs 76% respectively). In addition, there are significant differences in the known contraceptive methods (namely depo and male condom), and in the preferred contraceptive methods (depo, implants, male condom and traditional methods).

Comparing with the national data from the Ethiopian Demographic and Health Survey of 2016 (Central Statistical Agency & ICF International, 2016), more people in our sample are Protestant (63% vs 54% nationally). Further, our sample is in general higher educated. Nationally, 71% of women and 68% of men aging 20-24 have primary as their highest educational level which is much higher than our sample (25% of women and 31% of men in our sample). Moreover, 89% of women and 90% of men aging 20-24 did not complete secondary education in Ethiopia compared with 32% of women and 41% of men in this sample. At national level, 11% of women and 9% of men aging 20-24 attained a certain form of tertiary education which is lower than our sample where 23% of women and men started with any level of tertiary education. Additionally, it is striking that the current unemployment rate in the sample of 9% is much higher than the national rate of 3% for Ethiopians ranging 15-24 years old (The World Bank, 2020). Further, comparing with the national average of Ethiopians aging between 20-24 years, the percentage of married women is higher than men until the age of 34, after which the pattern reverses (Central Statistical Agency & ICF International, 2016). In this research with respondents aging 18-25 years, the opposite is found. The percentage of women reporting to be in a polygamous relationship (5%) lies in between the average for the SNNPR region (16%) and the national average (4%) for women aging 20-24 years old. Furthermore, the ideal FSP is comparable with the national data (4.5 children for women and 4.6 for men). The knowledge about contraception of women in the sample is lower compared to the national average of 99% for women (no information available for men).

Table 4-5: Socio-demographic characteristics of the respondents

Variable	Total		Female		Male		
	Mean	SE	Mean	SE	Mean	SE	
Number of observations	426		209		217		
Age (years)	20.74	2.31	19.89	1.98	21.55	2.33	***
Religion							
Orthodox (%)	33.57		30.62		36.41		
Catholic (%)	0.23		0.00		0.46		
Protestant (%)	63.62		67.94		59.45		*
Traditional religion (%)	1.17		0.48		1.84		
Other religion (Akale, Mekane eyesus and Hawariyat) (%)	1.4		0.96		1.84		
Residence							
Indigenous (father, grandfathers lived in this village) (%)	85.21		84.21		86.18		
Immigrated from another village in the same district (%)	10.56		11.00		10.14		
Immigrated from another district in the SNNPR region (%)	3.99		4.78		3.23		
Immigrated from another region in Ethiopia (%)	0.23		0.00		0.46		
Ethnic group							
Kusume (%)	1.41		0.48		2.30		
Wolaita (%)	1.17		0.48		1.84		
Gamo (%)	53.05		52.63		53.46		
Gofa (%)	0.70		0.96		0.46		
Zayse (%)	3.52		3.83		3.23		
Gidicho (%)	2.82		2.87		2.76		
Konso (%)	25.59		25.84		25.35		
Derashe (%)	11.50		12.44		10.60		
Kore (%)	0.23		0.48		0.00		
Education							
No Education (%)	3.52		4.31		2.76		
Primary schooling incomplete (%)	14.79		11.98		17.52		
Primary schooling complete (%)	9.62		8.61		10.60		
High school incomplete (%)	8.69		7.18		10.14		
High school complete (%)	32.86		37.32		28.57		*
Preparatory schooling incomplete (%)	1.41		0.48		2.30		
Preparatory schooling complete (%)	6.34		7.66		5.07		
TVET incomplete (%)	11.74		12.92		10.60		
TVET complete (%)	8.21		6.23		10.14		
University - bachelor's degree (%)	2.82		3.35		2.30		
Principal occupation							
Student (%)	36.38		39.23		33.64		
Farmer							
Family labourer on own farm (%)	21.83		12.91		30.42		***
Wage employed in agriculture (%)	0.47		0.00		0.92		
Work in household							
Family labourer in the household (%)	11.27		18.18		4.61		**
Non-farming related job							
Family labourer in a non-farm business (%)	0.94		1.44		0.46		
Own non-farm business (%)	13.85		11.96		15.67		
Hired servant (%)	1.88		0.48		3.23		**
Civil servant (%)	1.41		0.96		1.84		
Entrepreneur (%)	0.7		0.96		0.46		
Religious job (%)	0.47		0.48		0.46		
Charity worker (%)	1.88		2.87		0.92		
Looking for a job (%)	8.92		10.53		7.37		

*Note: Comparison between the male and female subgroups was based on: a) for ratio data: a two-sided t-test with post-hoc Tukey test (if the normality and homoscedasticity assumption is valid) or a Kruskal-Wallis test with post-hoc Dunn's test accounting for the False Discovery Rate (if the normality and homoscedasticity assumption do not hold) and b) for nominal data: a pairwise Chi Square test accounting for the False Discovery Rate. Significant differences of the mean of the male vs. female subgroups are shown with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

Table 4-6: Respondent's characteristics and preferences concerning marriage, child-raising and contraception

Variable		Total		Female		Male		
		Mean	SE	Mean	SE	SE	Mean	
Number of observations		426		209		217		
Marital status								
Married (%)		23.00		19.14		26.73		*
Age at first marriage (years)		19.12	2.42	17.83	1.93	20.02	2.32	***
Polygamous relationship (%)	Total	/		5.00		/		
	Mean number of wives	/		2[1]		/		
Not married (%)	Total	77.00		80.86		73.27		*
	Single	99.7		99.41		100.00		
	Divorced	0.3		0.59		0.00		
Preference to become married (%)	Total	85.06		86.98		83.02		
	Age	25.12	2.78	23.83	2.27	26.55	5.59	***
Currently having children								
Yes (%)		19.48		16.27		22.58		
Current number of children	Total number	1.74	1.02	1.71	1.13	1.76	0.95	
	Boys	0.89	0.69	0.94	0.68	0.86	0.71	
	Girls	0.86	0.84	0.77	0.84	0.92	0.84	
Having deceased children (%)		7.14		8.57		6.12		
	Total number	1.33	0.52	1.33	0.58	1.33	0.58	
	Boys	0.83	0.75	1.00	1.00	0.67	0.58	
	Girls	0.5	0.55	0.33	0.58	0.67	0.58	
Age at first birth (years)		19.71	2.24	18.66	2.10	20.47	2.03	***
No (%)		80.52		83.73		77.42		
Preferred age at first birth (years)		26.26	3.12	24.91	2.36	27.65	3.20	***
Ideal number of children	Total	4.32	1.20	4.29	1.11	4.35	1.28	
	Boys	2.28	0.74	2.25	0.66	2.30	0.82	
	Girls	2.04	0.70	2.03	0.71	2.05	0.70	
Ideal level of education	Boys Primary school (%)	0.00		0.00		0.00		
	High school (%)	0.23		0.00		0.46		
	Preparatory school (%)	2.11		0.95		3.53		
	TVET (%)	1.88		1.43		2.30		
	University (%)	95.77		97.61		94.01		*
	Girls Primary school (%)	0.23		0.00		0.46		
	High school (%)	2.58		1.44		3.69		
	Preparatory school (%)	2.58		2.39		2.76		
	TVET (%)	2.82		2.87		2.76		
	University (%)	91.78		93.30		90.32		
Contraception								
Knowledge about contraception (%)		82.39		76.08		88.48		***
Known methods (%)	Depo (injection e.g. jadelle)	24.77		31.60		19.12		***
	Contraceptive pills	9.40		11.25		7.88		
	Implants	23.16		25.33		21.36		
	Loop or intra-uterine device (IUD)	5.90		6.12		5.72		
	Male condom	19.63		9.78		27.78		***
	Female condom	2.81		2.94		2.71		
	Tubal/dicotomay (long-term method for women)	0.71		0.78		0.65		
	Vasectomy (long-term method for men)	0.54		0.58		0.51		
	Lactational amenorrhea method (LAM)	0.10		0.22		0.00		
	Post-pill	6.20		6.22		6.18		
	Traditional/natural methods	6.77		5.19		8.08		
Preferred methods (%)	Depo (injection e.g. jadelle)	28.77		40.25		19.27		***
	Contraceptive pills	1.42		2.52		0.52		
	Implants	34.19		47.17		23.44		***
	Loop or intra-uterine device (IUD)	3.99		4.40		3.65		
	Male condom	23.08		0.00		42.19		***
	Tubal/dicotomay (long-term method for women)	0.57		1.26		0.00		
	Vasectomy (long-term method for men)	0.28		0.00		0.52		
	Post-pill	0.85		1.26		0.52		
	Traditional/natural methods	6.84		3.14		9.90		**
Ever used contraception (%)		48.72		52.83		45.31		
Currently using contraception (%)		44.44		40.48		48.28		
Having access to contraception in this village								
Yes (%)	Total	95.44		95.60		95.31		
Shortage	Never	39.40		44.08		35.52		
	Few times a year	15.82		13.16		18.03		
	Few times a month	13.43		11.84		14.75		
	Always	0.30		0.66		0.00		
	Don't know	31.04		30.26		31.69		
No (%)	Total	4.48		4.35		4.59		
	Preference for having more access (%)	81.25		100.00		66.67		

[1] Both respondents are the second wife of their husband.

*Note: Comparison between the male and female subgroups was based on: a) for ratio data: an ANOVA test with post-hoc Tukey test (if the normality and homoscedasticity assumption is valid) or a Kruskal-Wallis test with post-hoc Dunn's test accounting for the False Discovery Rate (if the normality and homoscedasticity assumption do not hold) and b) for nominal data: a pairwise Chi Square test accounting for the False Discovery Rate. Significant differences of the mean of the male vs. female subgroups are shown with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

Comparing the ideal number of children among different subgroups of the respondents shows that respondents currently having children or being married prefer a significantly higher number of children, boys and girls than respondents without children (Table 4-7). Further, respondents from the Gamo Gofa Zone prefer a higher number of children and boys compared to respondents from the Segen People's Zone.

Table 4-7: Ideal number of children for different subgroups of the respondents

		Ideal number of children								
		Total			Boys			Girls		
		Mean (SE)			Mean (SE)			Mean (SE)		
Full sample		4.31			2.27			2.04		
Comparison group 1	Comparison group 2	1	2		1	2		1	2	
Female	Male	4.29	4.35		2.25	2.30		2.03	2.05	
		(0.08)	(0.09)		(0.05)	(0.06)		(0.05)	(0.05)	
Currently having children	Currently not having children	4.71	4.22	***	2.43	2.24	***	2.28	1.98	***
		(0.14)	(0.06)		(0.08)	(0.04)		(0.09)	(0.04)	
Orthodox	Protestant	4.34	4.31		2.29	2.27		2.04	2.04	
		(0.11)	(0.07)		(0.06)	(0.05)		(0.06)	(0.04)	
Indigenous	Immigrated	4.34	4.22		2.28	2.27		2.06	1.95	
		(0.06)	(0.14)		(0.04)	(0.08)		(0.04)	(0.09)	
No education	Secondary schooling or higher	4.20	4.17		2.27	2.22		1.93	1.94	
		(0.43)	(0.06)		(0.21)	(0.04)		(0.27)	(0.03)	
Student	No students	4.22	4.38		2.23	2.31		1.99	2.07	
		(0.09)	(0.07)		(0.06)	(0.05)		(0.05)	(0.05)	
Farmer	Non-farming job	4.33	4.32		2.27	2.28		2.06	2.03	
		(0.13)	(0.06)		(0.08)	(0.04)		(0.08)	(0.04)	
Married	Not married	4.65	4.22	***	2.43	2.23	***	2.22	1.98	***
		(0.13)	(0.06)		(0.08)	(0.04)		(0.08)	(0.04)	
Monogamous	Polygamous	4.47	4.50		2.34	2.50		2.13	2.00	
		(0.15)	(0.50)		(0.09)	(0.50)		(0.11)	(0.00)	
Knowledge about contraception	No knowledge about contraception	4.33	4.27		2.29	2.23		2.04	2.04	
		(0.07)	(0.12)		(0.04)	(0.07)		(0.04)	(0.07)	
Ever using contraception	Never using contraception vs. not	4.35	4.30		2.30	2.27		2.05	2.03	
		(0.09)	(0.08)		(0.06)	(0.05)		(0.05)	(0.05)	
Currently using contraception	Currently not using contraception	4.46	4.26		2.37	2.24		2.09	2.02	
		(0.14)	(0.06)		(0.08)	(0.04)		(0.09)	(0.04)	
Segen People's Zone	Gamo Gofa Zone	4.22	4.38	*	2.13	2.37	***	2.09	2.01	
		(0.09)	(0.07)		(0.05)	(0.05)		(0.06)	(0.04)	

*Note: Comparison between the subgroups was based on: a) for ratio data: a two-sided t-test (if the normality and homoscedasticity assumption is valid) or a Mann-Whitney test (if the normality and homoscedasticity assumption do not hold) and b) for nominal data: a pairwise Chi Square test accounting for the False Discovery Rate. Significant differences of the mean of the subgroups are shown with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

4.2.1. Self-expressed reasons for the heterogeneity in fertility preferences

Stated reasons for the ideal number of children among respondents with and without children are mainly economic (44% of the respondents), preference for quality of child-raising (25%) and to secure help in the household (10%) (Table 4-8). So, the respondents explicitly stated themselves to find the quality aspect of child-raising to be important.

Among the 118 respondents out of 426 who stated to have a higher preference for boys (obtained from Table 4-6), the main reasons are to secure help in the household (45% of the respondents) and influence from the community (39%). However, the main reasons for a preference for girls for 44 respondents out of 426 are to secure help in the household (52%). Women reported this reason significantly more frequently. The second most reported reason is the continuation of the family lineage in the village (27%) which was stated significantly more frequently by the male respondents.

Table 4-8: Stated influences for the ideal number of children (multiple answers were possible)

Reasons	Total			Respondents with a preference for boys			Respondents with a preference for girls			
	Total	Women	Men	Total	Women	Men	Total	Women	Men	
Number of observations	N = 426	N = 209	N = 217	N = 118	N = 54	N = 64	N = 44	N = 22	N = 22	
To secure help in the household (%)	10.30	8.42	12.11	44.75	43.20	46.06	51.89	75.00	28.77	***
Wealth status (%)	1.31	0.96	1.65	2.81	0.61	4.67	3.02	4.55	1.50	
Gender preference (certain girl/boy ratio) (%)	5.05	5.01	5.10	5.36	6.48	4.42	15.91	9.09	22.73	
Influence from the community (%)	4.26	5.01	3.53	38.69	42.59	35.41	2.27	0.00	4.55	
Continuation of the family lineage (%)	1.93	1.22	2.62	8.31	7.09	9.34	26.89	11.36	42.41	**
To secure caretaking when parents become older (%)	4.43	4.47	4.39							
Decision of spouse (%)	1.45	1.11	1.77							
Economic reason (%)	43.94	44.38	43.53							
Preference for quality of child-raising (%)	24.73	26.22	23.30							
Health of the mother (%)	2.33	2.90	1.78							

*Note: Comparison between the subgroups was based on a pairwise Chi Square test accounting for the False Discovery Rate. Significant differences of the mean of the subgroups are shown with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

4.2.2. Heterogeneity in fertility preferences

The parameter estimates of the 'average' respondent's preferences reveal that there are significant differences in the preferences between male and female respondents for 13 of the 23 parameters of the MXLM (Table 4-9). Moreover, the output reveals that all the schooling components for boys and girls are significant (i.e. at least one of the terms in which the schooling components occur is significant). Remark that all the parameter estimates need to be interpreted with respect to the base level FNoPrimary (i.e. the number of girls not attaining any schooling) which was omitted to make the choice model of the mixture estimable.

Table 4-9: Parameter estimates of the full MXLM of the first order multiplicative Scheffé model with a quadratic effect of the amount. The MXLM includes interaction effects of the dummy variable gender (Female = 1; Male= 0).

Variable	Posterior mean		Posterior SD	Lower 95% CI	Upper 95% CI
FPrimary	-45.94	*	20.30	-258.40	-186.34
FSecondary	120.60		21.95	-20.47	55.59
FTertiary	91.98	*	28.58	-136.00	-28.61
MNoPrimary	-77.16	*	38.36	-352.44	-235.63
MPrimary	240.55	*	38.33	146.77	297.16
MSecondary	-155.74	*	30.81	-203.67	-101.73
MTertiary	-60.08	*	38.14	-213.91	-8.53
FNoPrimary*NChildren	17.15	*	11.37	-160.11	-93.60
FPrimary*NChildren	66.25	*	7.15	17.68	50.31
FSecondary*NChildren	83.19	*	16.05	20.08	45.45
FTertiary*NChildren	13.90	*	8.74	42.89	70.99
MNoPrimary*NChildren	-75.06		13.80	-19.21	14.40
MPrimary*NChildren	79.22	*	9.80	-153.75	-96.80
MSecondary*NChildren	202.08		12.47	-4.71	16.67
MTertiary*NChildren	1.08	*	0.98	136.98	210.76
FNoPrimary*NChildren^2	-6.33	*	0.74	4.02	10.53
FPrimary*NChildren^2	-6.34	*	1.38	-4.24	-1.52
FSecondary*NChildren^2	0.24	*	0.60	-8.03	-5.07
FTertiary*NChildren^2	4.80	*	1.06	-7.36	-4.96
MNoPrimary*NChildren^2	-4.20		0.91	-1.61	0.70
MPrimary*NChildren^2	-14.07	*	1.24	2.47	7.07
MSecondary*NChildren^2	-28.25	*	4.35	-2.94	-0.26
MTertiary*NChildren^2	-0.58	*	0.39	-18.08	-11.83
Gender*FPrimary	-161.27		39.29	-80.65	69.21
Gender*FSecondary	-14.86		38.04	-37.10	133.06
Gender*FTertiary	-0.33	*	49.05	-178.01	-16.79
Gender*MNoPrimary	62.99	*	83.77	-383.66	-67.46
Gender*MPrimary	104.58		48.37	-79.21	65.23
Gender*MSecondary	-45.47		52.60	-179.64	63.59
Gender*MTertiary	-197.15	*	43.38	-192.58	-10.65
Gender*FNoPrimary*NChildren	73.25	*	19.26	-124.05	-77.12
Gender*FPrimary*NChildren	54.95		16.91	-19.71	18.24
Gender*FSecondary*NChildren	59.47	*	15.77	16.50	63.37
Gender*FTertiary*NChildren	-48.51	*	17.03	70.63	112.15
Gender*MNoPrimary*NChildren	-83.54		12.82	-30.24	47.23
Gender*MPrimary*NChildren	72.58		9.32	-44.01	6.13
Gender*MSecondary*NChildren	161.82	*	17.44	53.34	118.32
Gender*MTertiary*NChildren	-2.43	*	1.54	92.26	142.15
Gender*FNoPrimary*NChildren^2	-3.36	*	1.41	5.57	10.91
Gender*FPrimary*NChildren^2	-2.71		1.51	-1.70	1.57
Gender*FSecondary*NChildren^2	4.71	*	1.57	-6.98	-3.07
Gender*FTertiary*NChildren^2	7.42	*	1.53	-9.25	-5.40
Gender*MNoPrimary*NChildren^2	-5.90		1.21	-4.44	2.25
Gender*MPrimary*NChildren^2	-10.29	*	1.47	-6.67	0.14
Gender*MSecondary*NChildren^2	-39.26	*	5.23	-12.98	-7.76
Gender*MTertiary*NChildren^2	-0.31	*	0.61	-13.91	-8.71
AICc			89.72		
BIC			226.91		
-2LL			47.54		

*Note: Significant differences of the posterior mean on a 5% level are indicated with *. The posterior SD refers to the posterior standard deviation. The explanation of the variables can be found in Table 3-2.*

Figure 4-1 (based on a first order multiplicative Scheffé model with a quadratic effect of the amount) indeed fulfils the **first assumption** of a concave relationship between the utility and the number of children. Moreover, it shows that men's utility for the number of children is higher than women's utility.

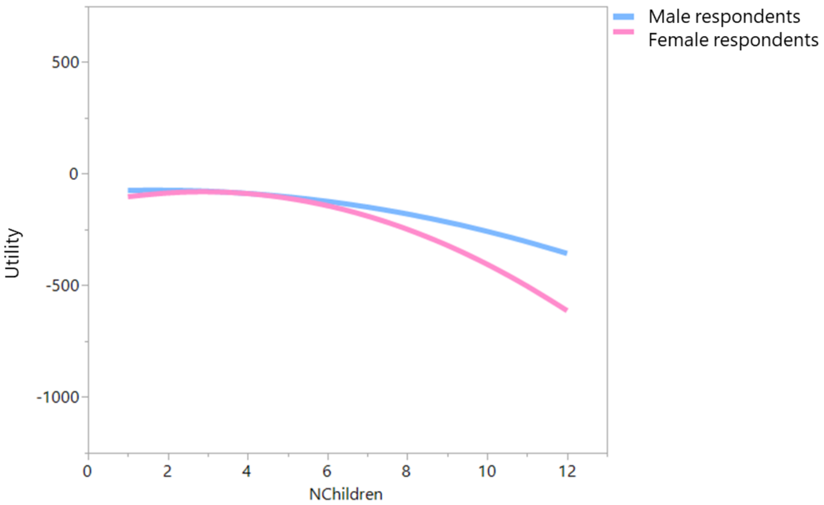


Figure 4-1: The respondent's utility in function of the number of children for male and female respondents respectively

The **second assumption** on which the selection of the utility function was based, is the use of interaction terms in the model between the number of children and the schooling components since this seemed to be more realistic. The MXLM parameter estimates are visualised in Figures 4-2 and 4-3 for the male and female respondents respectively. Visually, the results of male and female respondents are quite similar. For most of the graphs, at any given schooling attribute percentage, the utility is lower if the number of children increases. This reveals that the respondents' utility decreases if the number of children reaches high numbers. Further, the utility follows a decreasing course if the percentage FNoPrimary and MNoPrimary increases, whereas the utility increases if FTertiary and mainly MTertiary increases, with the highest utility for 4 to 8 children. This shows the high preference for tertiary schooling and the lower preference for no schooling, as expected. For the other schooling components (FPrimary, MPrimary, FSecondary and MSecondary), the slope of the utility increases or decreases dependent on the number of children. Note also that overall, the utility is almost negative for FNoPrimary, FPrimary, MNoPrimary, MPrimary and MSecondary while it shows positive values for the attributes FSecondary, FTertiary and MTertiary.

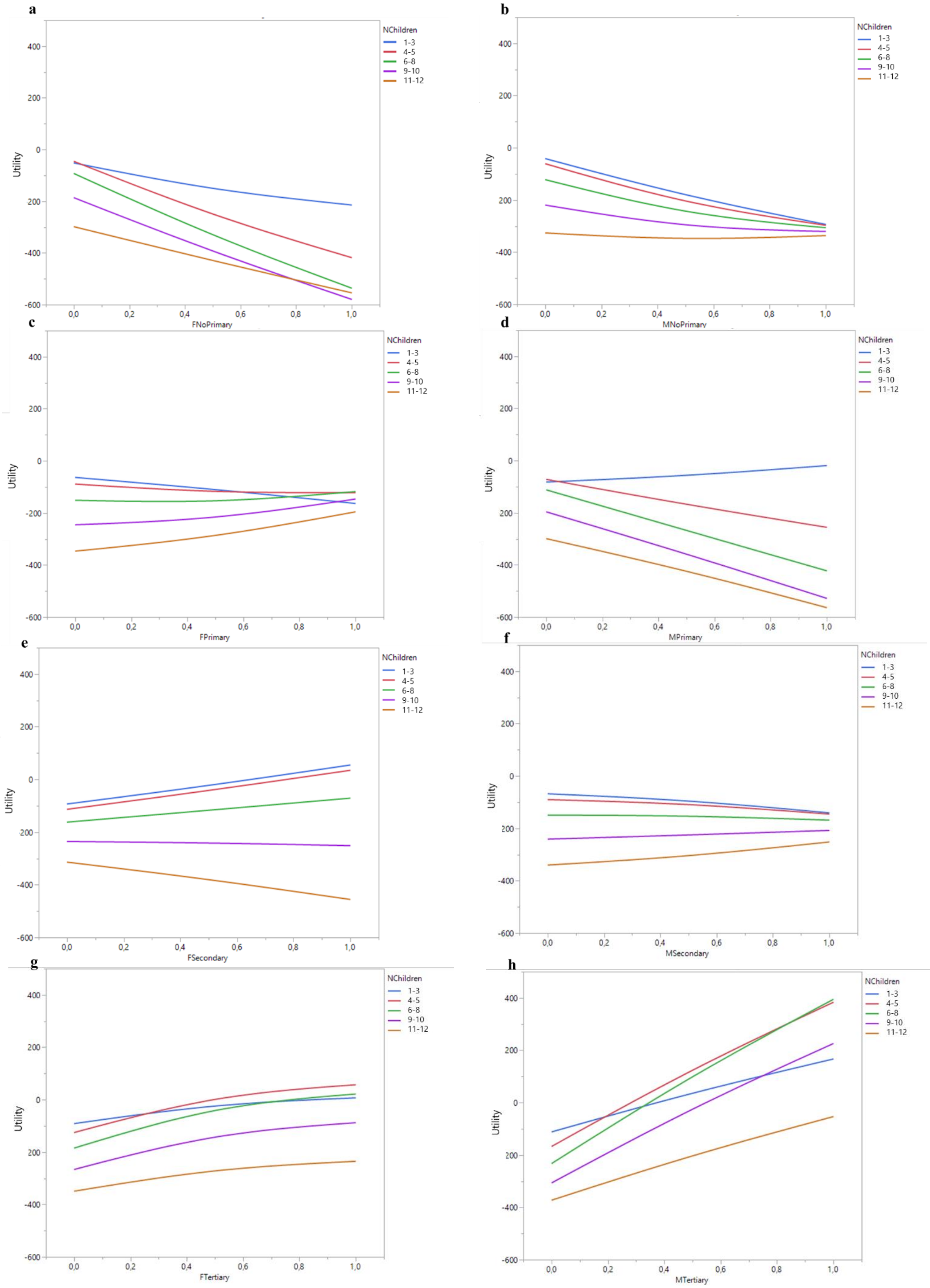


Figure 4-2: Visualization of the MXLM parameter estimates of ***male respondents****' utility in function of the different schooling components for different ranges of the number of children. Note: all results need to be interpreted with respect to FNoPrimary.*

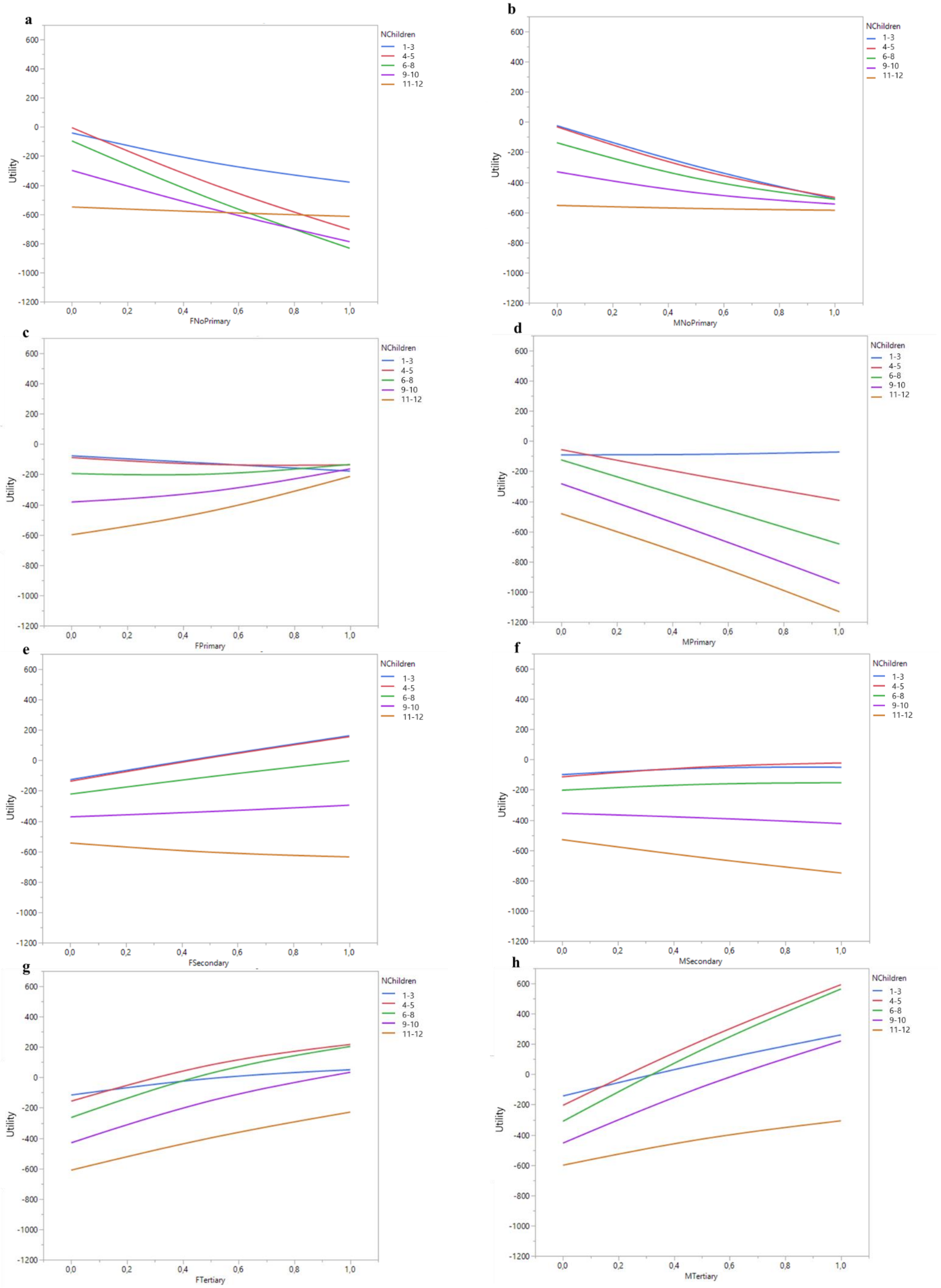


Figure 4-3: Visualization of the MXLM parameter estimates of ***female respondents****' utility in function of the different schooling components for different ranges of the number of children. Note: all results need to be interpreted with respect to FNoPrimary.*

4.2.3. Optimal mixtures

From an optimization perspective, Table 4-10 shows the optimal mixture compositions. These are visualized in Figure 4-4.

Table 4-10: Optimal mixtures of the schooling components at the highest desirable number of children for the male and female respondents for the MXLM

Attributes		Male	Female
Number of children	Total	5.98	5.62
	Boys	3.17	2.75
	Girls	2.81	2.87
FNoPrimary		0.00	0.00
FPrimary		0.07	0.08
FSecondary		0.14	0.21
FTertiary		0.26	0.22
MNoPrimary		0.06	0.03
MPrimary		0.06	0.03
MSecondary		0.06	0.11
MTertiary		0.35	0.32

Note: No standard deviations could be obtained via the JMP Pro 14 software. The explanation of the variables can be found in Table 3-2.

The male and female respondents' maximized utility is obtained around 5.98 and 5.62 children respectively according to the DCE. Moreover, compared to the base level of FNoPrimary, men's utility is maximized if most of their children are boys who are able to finish tertiary schooling (35%) and girls who finish tertiary schooling (26%). Women's utility has an optimum if most of their children are boys and girls completing tertiary schooling (32% and 22% respectively) or girls who are able to finish secondary schooling (22% of their total number children).

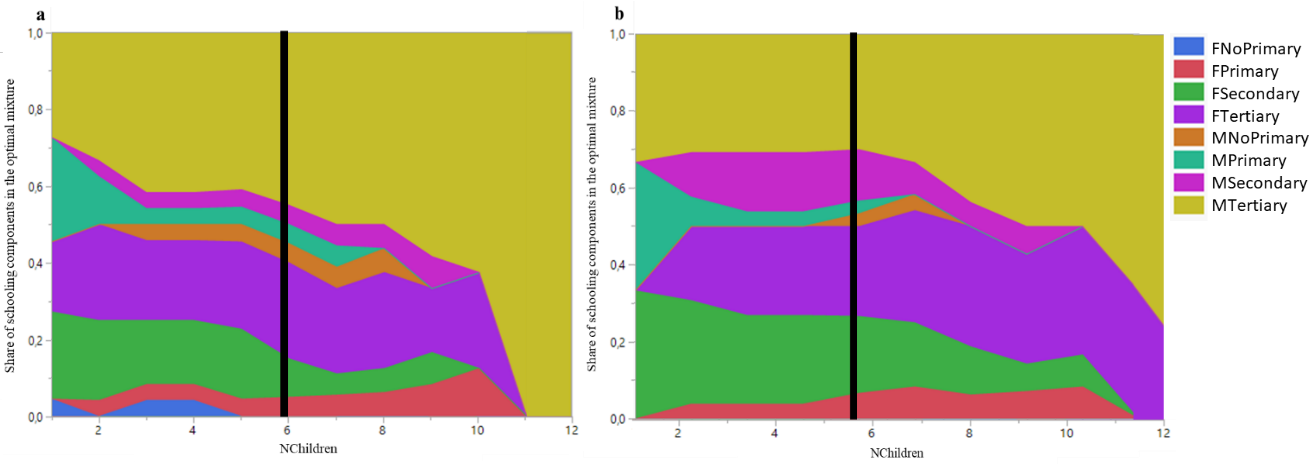


Figure 4-4: Optimal mixtures for a) the male and b) the female respondents from the MXLM for the schooling components as a function of the highest desirable number of children (indicated by the black line). Note: all results need to be interpreted with respect to FNoPrimary. The explanation of the variables can be found in Table 3-2.

4.2.4. The quantity-quality trade-off

The MRS for having one child more in terms of a percentage increase or decrease of the schooling components is shown along the indifference curves (i.e. constant-utility curves) in Figure 4-5 and 4-6 for the MXLM. The brightest red lines indicate the highest utility levels. These lines reveal the trade-off respondents make when they maximize their utility. A decreasing slope indicates that respondents are willing to trade off schooling and number of children, whereas an increasing slope indicates that respondents are only willing to accept one more child if more of their children are able to attain the corresponing educational level. Furthermore, the steepness of the slope reveals whether the respondents are willing to trade-off (in case of a decreasing slope) or gain (in case of an increasing slope) a large or relatively small amount of the quality component for one child extra.

The analysis of the marginal rate of subsitution reveals that both male and female respondents want more children if less of their daughters do not attain any educational level (i.e. FNoPrimary). The same holds for the sons (i.e. MNoPrimary) of male respondents. Female respondents are willing to trade-off schooling and their number of sons (i.e. MNoPrimary) until they have around three children. However, from circa three children onwards, they are only willing to have more children if more of their sons do attain any educational level. For the primary, secondary and tertiary schooling components, the trade-off also mostly applies only up to a certain number of children.

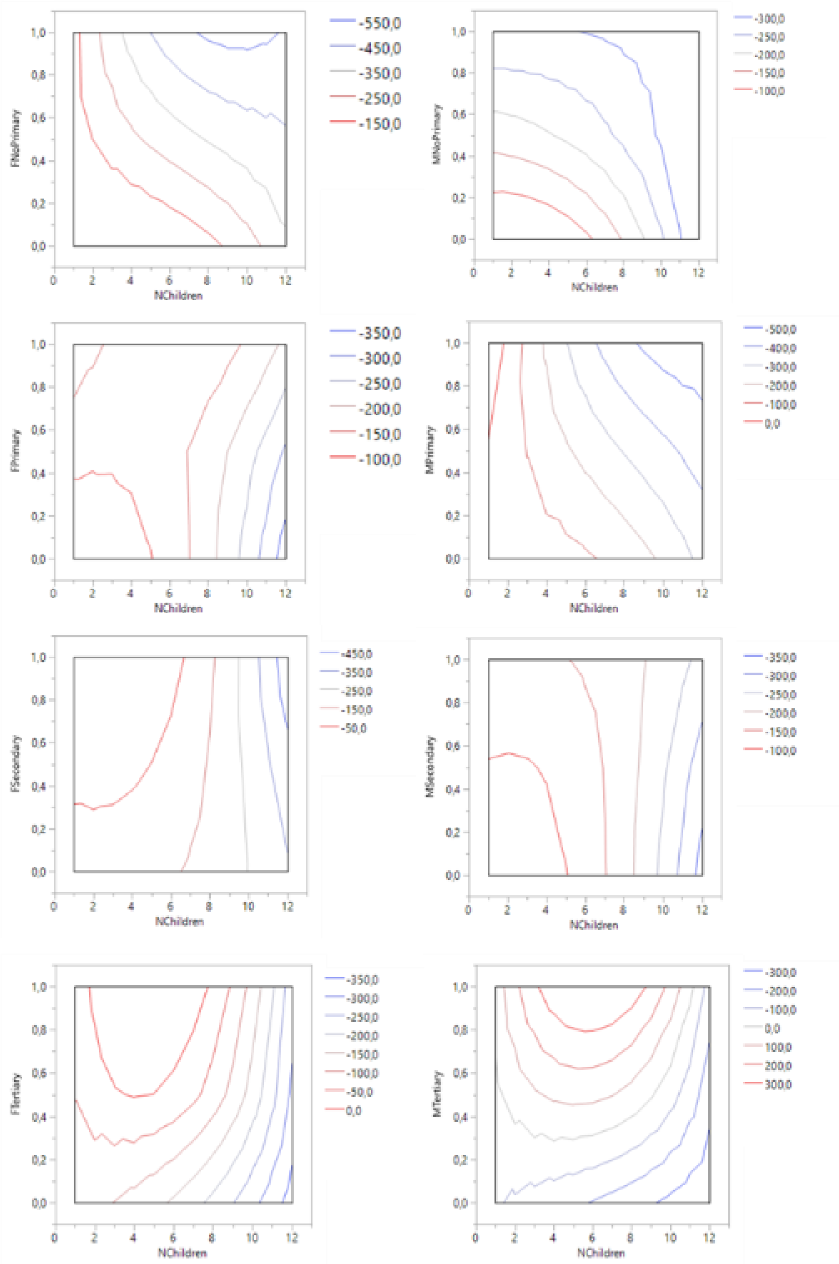


Figure 4-5: Utility indifference curves for the different schooling components (y-axis) and the number of children (x-axis) to visualize the MRS for the ***male*** *respondents of the MXLM. The brightest red lines show the highest constant-utility curves. Note: all results need to be interpreted with respect to FNoPrimary. The explanation of the variables can be found in Table 3-2.*

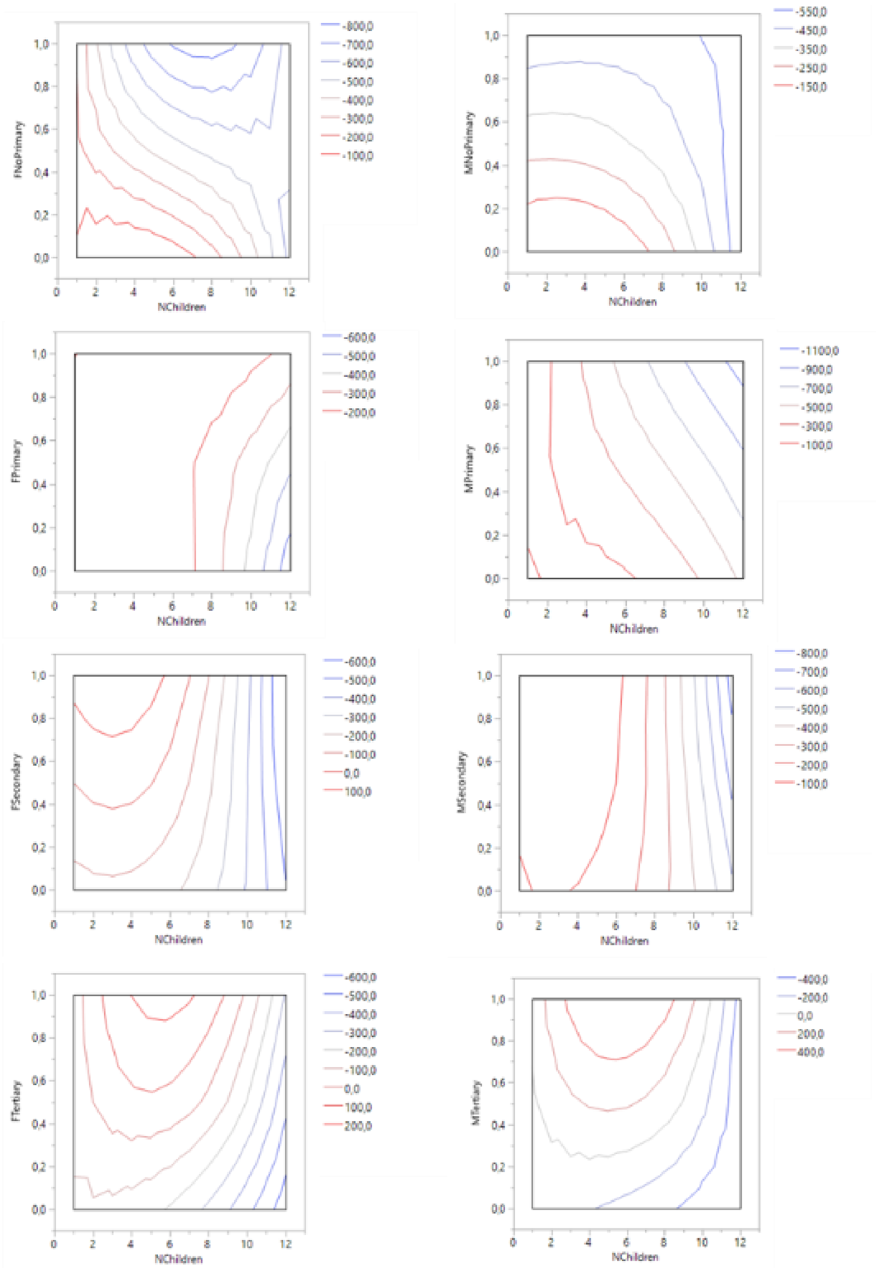


*Figure 4-6: Utility indifference curves for the different schooling components (y-axis) and the number of children (x-axis) to visualize the MRS for the **female** respondents of the MXLM. The brightest red lines show the highest constant-utility curves. Note: all results need to be interpreted with respect to FNoPrimary. The explanation of the variables can be found in Table 3-2.*

4.3. Fertility preferences beyond gender (LCM)

4.3.1. Sociodemographic characteristics

Table 4-11 shows the differences in sociodemographic characteristics between the five LCs. The results show that the LCs differ with respect to their zone and their ethnic group. Furthermore, the groups have a significantly different preference for their ideal number of children. However, they do not have a significantly different choice for the ideal educational level for their children.

Table 4-11: Differences in sociodemographic characteristics and preferred ideal number of children and level of education between the five latent classes (LCs) obtained during the follow-up questions

Variable	LC I			LC II			LC III			LC IV			LC IV		
	Mean	SE		Mean	SE		Mean	SE		Mean	SE		Mean	SE	
Number of observations	N = 71			N = 85			N = 94			N = 53			N = 123		
Age (years)	20.75	2.45		20.62	2.20		20.36	2.13		20.83	2.56		21.07	2.32	
Sex (1 = female; 0 = male)	50.27			47.06			56.38			50.94			43.09		
Zone (1 = Gamo Gofa Zone; 0 = Segen People's Zone)	35.21			50.59		‡	42.55			32.08			31.71		¥
District: Konso (%)	22.53			38.82		‡	25.53			20.75			20.33		¥
Indigenous (%)	77.46			89.41			85.11			79.25			89.43		
Religion: Protestant (%)	69.01			65.88			56.38			58.49			66.67		
Religion: Orthodox (%)	29.58			32.94			39.36			35.85			30.89		
Ethnic group: Gamo (%)	45.07			48.24			52.13			62.26			57.72		
Ethnic group: Konso (%)	23.94			38.82		‡‡	25.53			20.75			19.51		¥¥
Ethnic group: Derashe (%)	11.27			10.59			15.96			7.55			10.57		
Primary education (%)	28.17			22.35			21.28			28.30			24.39		
High school (%)	36.62			45.88			42.55			37.74			42.48		
Preparatory school (%)	11.27			3.53			9.57			1.89			9.75		
TVET (%)	18.31			23.53			19.15			26.42			16.26		
University (%)	2.82			2.35			2.13			0.00			4.85		
Years of schooling	10.03	3.36		10.11	3.03		9.79	3.64		9.77	3.48		10.22	3.29	
Student (%)	36.62			36.47			39.36			28.30			37.40		
Farmer (%)	22.54			24.71			21.28			24.53			21.95		
Married (%)	21.13			20.00			22.34			26.42			25.20		
Polygamous (%)	7.04			8.24			11.70			11.32			7.32		
Currently having children (%)	15.49			12.94			21.28			24.53			22.76		
Having deceased children (%)	4.23			0.00			0.00			1.89			1.63		
(Preferred) age at first birth (years)	25.61	3.86		24.96	3.04		24.87	4.30		24.60	3.97		24.84	4.27	
Knowledge about contraception (%)	77.46			88.24			73.40			84.91			86.99		
Ever used contraception (%)	32.39			44.71			30.85			39.62			48.78		
Access to contraception (%)	71.83			85.88			68.09			81.13			84.55		
Ideal number of children															
Total	4.52	1.61	¥¥	3.99	0.81	**¤¤‡‡‡	4.29	1.28		4.42	0.80	¥¥	4.41	1.19	¥¥¥
Boys	2.41	0.99		2.12	0.47		2.24	0.71		2.30	0.61		2.33	0.79	
Girls	2.11	0.80		1.87	0.55		2.04	0.79		2.11	0.54		2.08	0.72	
Ideal level of education: boys															
Primary school (%)	0.00			0.00			0.00			0.00			0.00		
High school (%)	0.00			0.00			0.00			0.00			0.81		
Preparatory school (%)	1.41			3.53			1.06			0.00			3.25		
TVET (%)	2.82			1.18			3.19			0.00			1.63		
University (%)	95.77			95.29			95.74			100.00			94.31		
Ideal level of education: girls															
Primary school (%)	0.00			0.00			0.00			0.00			0.81		
High school (%)	1.41			4.71			1.06			1.87			3.25		
Preparatory school (%)	4.23			2.35			3.19			1.89			1.63		
TVET (%)	5.63			3.53			3.19			1.89			0.81		
University (%)	88.73			89.41			92.55			94.34			93.50		

*Note: Comparison between LCs was based on: a) for ratio data: an ANOVA test with post-hoc Tukey test (if the normality and homoscedasticity assumption was valid) or the Kruskal-Wallis test with post-hoc Dunn's test accounting for the False Discovery Rate (if the normality and homoscedasticity assumption do not hold) and b) for nominal data: pairwise Chi Square test accounting for the False Discovery Rate. Significant differences from group 1 are indicated with * p<0.1, ** p<0.05, *** p<0.01, significant differences from group 2 with ¥ p<0.1, ¥¥ p<0.05, ¥¥¥ p<0.01, significant differences from group 3 with ¢ p<0.1, ¢¢ p<0.05, ¢¢¢ p<0.01, significant differences from group 4 with ¤ p<0.1, ¤¤ p<0.05, ¤¤¤ p<0.01 and significant differences from group 5 with ‡ p<0.1, ‡‡ p<0.05, ‡‡‡ p<0.01.*

4.3.1. Self-expressed reasons for the heterogeneity in fertility preferences

No significant differences in the reasons for the preferred number of children are found (Table 4-12). The main reasons for the preferred number of children are economic followed by a preference for quality of child-raising. Further, no significant differences are found between the LCs for the reasons of the respondents with a higher preference for boys or girls probably due to the small number of respondents in these classes. For all the classes, the main reasons for a preference for boys are to secure help in the household and influence from the community. Furthermore, the main reasons that some respondents have a higher preference for girls are to secure help in the household and the continuation of the family lineage.

Table 4-12: Self-expressed influences for the ideal number of children for the latent classes (LCs)

Reasons	LC I	LC II	LC III	LC IV	LC V
Reasons for the self-expressed ideal number of children					
Number of observation	N = 71	N = 85	N = 94	N = 53	N = 123
To secure help in the household (%)	16.07	6.21	9.54	9.92	10.72
To secure caretaking when parents become older (%)	5.52	1.51	4.22	7.72	4.56
Wealth status (%)	0.00	1.18	0.63	0.00	3.25
Decision of spouse (%)	1.52	0.62	1.41	1.89	1.81
Gender preference (certain girl/boy ratio) (%)	2.92	6.46	5.19	7.51	4.15
Influence from the community (%)	4.93	1.76	3.01	5.32	6.09
Economic reason (%)	39.28	50.20	39.01	46.08	45.16
Preference for quality of child-raising (%)	24.99	27.36	32.28	19.36	19.33
Health of the mother (%)	2.10	3.26	2.11	0.00	3.00
Continuation of the family lineage (%)	2.48	1.21	2.37	1.94	1.77
Reasons for a preference for boys					
Number of observations	N = 18	N = 24	N = 24	N = 16	N = 36
To secure help in the household (%)	51.83	49.29	32.58	50.00	43.97
Wealth status (%)	2.78	5.54	1.38	0.00	1.39
Gender preference (certain girl/boy ratio) (%)	10.17	4.17	8.33	6.25	1.39
Influence from the community (%)	27.78	37.50	46.50	37.50	40.28
Continuation of the family lineage in this village (%)	7.39	3.46	9.67	6.25	12.03
Reasons for a preference for girls					
Number of observations	N = 11	N = 9	N = 9	N = 4	N = 11
To secure help in the household (%)	60.00	25.00	45.00	66.63	63.33
Wealth status (%)	0.00	16.67	0.00	4.13	0.00
Gender preference (certain girl/boy ratio) (%)	0.00	33.33	20.00	6.25	13.33
Influence from the community (%)	0.00	0.00	0.00	12.50	0.00
Continuation of the family lineage in this village (%)	40.00	25.00	35.00	6.25	23.33

*Note: Comparison between LCs was based on the pairwise Chi Square test accounting for the False Discovery Rate. Significant differences from group 1 are indicated with * p<0.1, ** p<0.05, *** p<0.01, significant differences from group 2 with ¥ p<0.1, ¥¥ p<0.05, ¥¥¥ p<0.01, significant differences from group 3 with ¢ p<0.1, ¢¢ p<0.05, ¢¢¢ p<0.01, significant differences from group 4 with ¤ p<0.1, ¤¤ p<0.05, ¤¤¤ p<0.01 and significant differences from group 5 with ‡ p<0.1, ‡‡ p<0.05, ‡‡‡ p<0.01.*

4.3.2. Heterogeneity in fertility preferences

Preferences differ between respondents. Table 4-13 shows the five groups of respondents which have similar preferences. The visualization of the parameter estimates can be found in Appendix (Figures 8-5 and 8-6 to 8-11). Figure 8-5 reveals the existence of a concave relationship between the utility and the number of children for all the LC's (cf. the first assumption of the model). Figures 8-6 to 8-11 show that LCM I has a slightly higher preference for boys and mainly prefer tertiary schooling for boys and girls or secondary schooling for boys. LC II's utility increases most for boys completing secondary or tertiary schooling. Next, LC III's preference is highest for boys finishing tertiary schooling followed by girls completing tertiary schooling. Further, LC IV clearly has a preference for girls completing tertiary schooling followed by boys attaining this educational level. Finally, LC V seems to have the highest preference for boys finishing tertiary schooling.

Table 4-13: Parameter estimates of the five classes of the LCM of the first order multiplicative Scheffé model with a quadratic effect of the amount.

Term	LC I				LC II				LC III				LC IV				LC V			
	Estimate		Lower 95% CI	Upper 95% CI	Estimate		Lower 95% CI	Upper 95% CI	Estimate		Lower 95% CI	Upper 95% CI	Estimate		Lower 95% CI	Upper 95% CI	Estimate		Lower 95% CI	Upper 95% CI
FPrimary	7.27	*	1.62	12.92	262.91		56.87	468.96	-16.84	*	-31.60	-2.07	-17.57	*	-37.89	2.75	79.80	*	41.13	118.48
FSecondary	10.50	*	4.86	16.15	263.01	*	49.33	476.68	-29.60	*	-41.39	-17.81	-13.73		-35.91	8.45	107.31	*	65.10	149.51
FTertiary	9.25	*	3.06	15.45	183.54	*	54.24	312.85	-11.17		-26.75	4.41	-24.59	*	-49.78	0.60	101.92	*	57.34	146.50
MNoPrimary	12.35	*	4.54	20.16	352.86	*	79.58	626.15	-45.09	*	-69.65	-20.54	-0.77		-30.56	29.02	75.98	*	33.85	118.12
MPrimary	7.48	*	1.44	13.53	250.47	*	62.49	438.45	12.04	*	-1.95	26.03	-9.88		-30.70	10.93	166.31	*	99.02	233.60
MSecondary	7.62	*	2.36	12.88	260.07	*	62.06	458.07	-6.18		-22.76	10.39	-14.77		-32.92	3.38	70.92	*	36.12	105.72
MTertiary	8.98	*	3.17	14.79	303.03	*	68.01	538.05	-8.87		-22.61	4.87	-11.89		-34.81	11.03	80.62	*	40.19	121.05
FNoPrimary*NChildren	2.11	*	0.38	3.84	79.84	*	18.51	141.17	-9.22	*	-13.89	-4.54	-6.86	*	-12.81	-0.92	19.64	*	8.12	31.17
FPrimary*NChildren	0.51		-0.15	1.18	16.36	*	3.54	29.18	0.81		-0.63	2.25	2.45	*	-0.31	5.20	3.19	*	0.24	6.15
FSecondary*NChildren	0.00		-0.62	0.61	14.71	*	3.75	25.67	6.39	*	2.90	9.87	2.89		-1.54	7.33	-4.83	*	-7.17	-2.49
FTertiary*NChildren	0.18		-0.64	1.00	34.93	*	1.41	68.45	1.38		-0.45	3.22	12.49	*	3.63	21.34	6.41	*	2.82	9.99
MNoPrimary*NChildren	-1.18		-2.65	0.29	-23.47	*	-41.08	-5.87	7.95	*	3.92	11.99	-9.69	*	-19.33	-0.04	6.86	*	2.21	11.51
MPrimary*NChildren	-0.31		-1.19	0.57	14.35	*	-1.54	30.24	-7.73	*	-10.61	-4.86	-2.14		-5.43	1.16	-29.01	*	-41.53	-16.49
MSecondary*NChildren	0.78	*	0.10	1.47	1.44		-1.41	4.30	-2.18	*	-4.50	0.14	1.85		-1.69	5.38	7.08	*	3.84	10.33
MTertiary*NChildren	0.83	*	0.19	1.47	-10.96	*	-21.28	-0.63	3.02	*	1.02	5.02	6.15	*	2.53	9.77	15.36	*	9.49	21.23
FNoPrimary*NChildren*NChildren	-0.14	*	-0.26	-0.02	-5.70	*	-10.07	-1.33	0.69	*	0.39	1.00	0.51	*	0.06	0.97	-0.76	*	-1.52	0.00
FPrimary*NChildren*NChildren	-0.04		-0.09	0.02	-1.83	*	-3.34	-0.33	-0.07		-0.18	0.05	-0.16		-0.40	0.08	-0.36	*	-0.64	-0.09
FSecondary*NChildren*NChildren	-0.02		-0.07	0.03	-1.44	*	-2.50	-0.38	-0.44	*	-0.67	-0.22	-0.29	*	-0.62	0.04	0.11		-0.03	0.26
FTertiary*NChildren*NChildren	0.00		-0.07	0.06	-2.76	*	-5.45	-0.07	-0.14	*	-0.27	0.00	-0.94	*	-1.59	-0.29	-1.26	*	-1.70	-0.81
MNoPrimary*NChildren*NChildren	0.06		-0.04	0.16	1.00	*	0.09	1.92	-0.49	*	-0.73	-0.24	0.75	*	0.02	1.48	-0.15		-0.47	0.18
MPrimary*NChildren*NChildren	0.04		-0.03	0.10	-1.40	*	-2.84	0.04	0.44	*	0.26	0.63	0.15		-0.12	0.42	1.42	*	0.61	2.24
MSecondary*NChildren*NChildren	-0.06	*	-0.11	0.00	0.37	*	0.00	0.73	0.10		-0.05	0.26	-0.13		-0.41	0.16	-0.76	*	-1.06	-0.46
MTertiary*NChildren*NChildren	-0.05	*	-0.10	0.00	1.33	*	0.12	2.54	-0.32	*	-0.47	-0.16	-0.41	*	-0.67	-0.15	-1.34	*	-1.89	-0.78
Latent class probability (%)	0.19				0.20				0.23				0.11				0.27			
AICc									4327.69											
BIC									5100.40											
-2LL									4083.98											

*Note: Significant differences of the mean of the posterior mean on a 5% level are indicated with *. The explanation of the variables can be found in Table 3-2.*

4.3.3. Optimal mixtures

The maximized utility for the different LCs occurs at 6.90, 6.00, 2.00, 7.30 and 4.00 children for LC I to V respectively (Table 4-14). The corresponding optimal schooling components are visualized in Figure 8-12 in Appendix. LC I, II, III and V have a higher preference for boys whereas only LC IV has a higher preference for girls. LC I's utility reaches a maximum if a large share of their children are able to finish tertiary schooling (39% boys and 28% girls) or if their boys can complete secondary schooling (20%). LC II's utility is optimized if most of their children can complete tertiary schooling (23% boys and 20% girls) followed by boys who are able to graduate from secondary (16%) or primary (17%) schooling and girls who do not finish any educational level (15%). LC III prefers that most of their boys are able to complete primary schooling (29%) or that most of their boys (22%) and girls (22%) can finish tertiary schooling. For LC IV, the share of the schooling components in their optimal mixture consists mainly of girls finishing secondary schooling (38%) followed by boys completing tertiary schooling (27%). Next, LC IV's optimum occurs if most of their children are able to finish tertiary schooling (44% boys and 27% girls).

Table 4-14: Optimal mixtures of the schooling components at the highest desirable number of children for the male and female respondents for the five LCs

Attributes		LC I	LC II	LC III	LC IV	LC V
Number of children	Total	6.90	6.00	2.00	7.30	4.00
	Boys	4.14	3.36	1.28	2.92	2.44
	Girls	2.76	2.64	0.72	4.38	1.56
FNoPrimary		0.00	0.15	0.05	0.04	0.00
FPrimary		0.07	0.08	0.07	0.09	0.06
FSecondary		0.05	0.04	0.02	0.09	0.06
FTertiary		0.28	0.17	0.22	0.38	0.27
MNoPrimary		0.00	0.00	0.01	0.00	0.12
MPrimary		0.01	0.17	0.29	0.04	0.01
MSecondary		0.20	0.16	0.12	0.09	0.04
MTertiary		0.39	0.23	0.22	0.27	0.44

Note: No standard deviations could be obtained via the JMP Pro 14 software. The explanation of the variables can be found in Table 3-2.

4.3.4. The quantity-quality trade-off

The visualization of the trade-off is shown in Figure 8-13 in Appendix for the LC's. For the five classes, just like for the MXLM, the trade-off of the schooling components for number of children also applies mostly up to a certain number of children. That number is situated around the optimal number of children for LCM I, II, IV and V but deviates from it for LCM III. Moreover, it is striking that LCM I, II and III make a real quantity-quality trade-off for the largest component of their optimal mixture, namely FSecondary and FTertiary for LCM I, MSecondary and MTertiary for LCM II and MPrimary for LCM III.

5. Discussion

In general, the results of the DCE described above show that parents derive utility from all the considered attributes. So, the quantity and gender of children, and quality of child-raising are important in fertility preferences. This is consistent with our modified model of Galor (2012) and de la Croix and Perrin (2016) described in 2.5.1. In addition, we find that respondents mainly made their choices based on gender and educational level. This implies that respondents have a high preference for schooling and a gender bias in preference for quantity and quality. In the following paragraphs, the results of the quantity, quality and gender aspect are discussed separately.

<u>Quantity attribute</u>

Our results related to the first research question of the heterogeneity in fertility preferences reveal different elements. Concerning the **general fertility preferences**, the DCE's optimal mixtures for the MXLM show that the stated (i.e. latent) male and female respondents' utility is maximized at 5.98 and 5.62 children respectively (Table 4-10). This is slightly higher than the current national TFR of 5.2 in rural areas (Central Statistical Agency & ICF International, 2016). Further, this stated FSP is 1.63 children for male and 1.33 children for female respondents higher than their self-expressed FSP during the follow-up questions. This shows the usefulness of our DCE method which is able to reveal latent fertility preferences. The discrepancy elicits two elements.

First, the higher observed stated preference for the number of children compared to the national average shows that it is possible that national data is biased. This is plausible because national data is based on survey data. Birth registration only became obligatory since 2012 (UNICEF, 2019), although currently the completeness of birth registration[28] is only 2.7% (The World Bank, 2018).

Second, the gap between stated/actual and self-expressed FSP is also observed in the literature, ranging from one to two children (Bongaarts & Casterline, 2013; Central Statistical Agency & ICF International, 2016; Günther & Harttgen, 2016; Pritchett, 1994; Trinitapoli & Yeatman, 2018; Van Lith et al., 2013). In literature, two opposing explanations are given. A first explanation is that the self-expressed FSP may be lower due to the hypothetical and research setting (i.e. hypothetical and social desirability bias) (Bongaarts & Casterline, 2013; Pritchett,

[28] Measured as the percentage of children under the age of five whose births were registered (The World Bank, 2020).

1994). Accordingly, the gap between stated/actual and self-expressed FSP can be explained by a higher prevailing FSP per se. A second explanation is an unmet need for family planning services (and hence family planning efforts) (Günther & Harttgen, 2016). The reason why this gap probably cannot be attributed to an unmet need for family planning services is threefold. Firstly, in this research 95% of the respondents reported having access to contraception in their village. Secondly, recent evidence reveals high levels of unrealized fertility in SSA (Casterline & Han, 2017). Thirdly, there is mixed evidence about the relationship between family planning policies focussing on contraception and fertility reduction in literature. Some authors attribute the steep decline in TFR in Eastern Africa to their family planning policies (Bongaarts, 2011; Gerland et al., 2014; Van Lith et al., 2013) whereas others dispute this (Galor, 2012; Günther & Harttgen, 2016; Pritchett, 1994). In Ethiopia, although family planning policies that focussed on improved access to contraception could have facilitated the changes, looking at the decline in TFR over time, the decreasing course started already before the implementation of the influential family planning policy of 1993. Olson and Piller (2013) report that the scarcity of natural and economic resources in Ethiopia already rose the recognition of the need for family planning among the population before the government intervened during history. These include the historic famines (1983-1985, 1998, 2006, and 2008), the water scarcity in 2002, the high rate of deforestation around 1993 together with the twofold increase in population from 1980 to 2010 resulting in average farm sizes of 0.8 ha (Food and Agricultural Organization, 2018) due to the system of shared inheritances by multiple sons (Office of the Prime Minister [Ethiopia], 1993; Population Reference Bureau, 2002). Therefore, family planning policies should mainly focus on decreasing the prevailing family size norm instead of focussing on contraception. Our results help to explain the low demand for contraception (i.e. contraceptive prevalence rate[29]) in Ethiopia (40%) (The World Bank, 2020) and suggest that the demand for contraceptives will follow once the desired FSP decreases. This can be done for example through mass media, health centres, education, etc.

Based on **anecdotal evidence** with some key informants from the national and regional government and the Arba Minch University, it was also found that improvements in the current family planning policies and programs must include broadening the method mix beyond the dominance of injections, improving the weak government control over budgets, lowering the dependence on donor support, decreasing the high turnover of health professionals and increasing evidence-based planning and birth registration (personal communication, 2019).

[29] Measured as the percentage of men and women married or in union between 15 and 49 years old practicing any form of contraception obtained via surveys (The World Bank, 2020).

Concerning the **heterogeneity in fertility preferences**, our results elicit three elements. First, the follow-up questions show that a higher self-expressed FSP is related with *being married*, *having children* and *place of residence* (i.e. living in the Gamo Gofa Zone) (Table 4-7). Second, no heterogeneity is found concerning the reasons given by respondents for their self-expressed FSP (Table 4-8 and 4-12). Third, the DCE reveals that a higher stated FSP is associated with *being male*, *place of residence* (i.e. living in the Gamo Gofa Zone or Konso district), *belonging to the Konso ethnic group* and the *self-expressed FSP* (Table 4-9 and 4-11). It is remarkable that the optimal number of children based on the DCE differs in relative magnitude from the self-expressed FSP for the LCs. This can be explained by the difference in method. The optimal number of children is obtained via the maximized desirability of the utility function whereas the self-expressed FSP is acquired via simple averages.

Overall, the main determinants explaining the heterogeneity in fertility preferences in our study area are, as classified according to Bongaarts (1978, 2015, 2017), the *self-expressed FSP*, the direct factor *marital status* and indirect factors related to the cultural diffusion theory (*gender* and *ethnicity*) and economic and investment theory (*place of residence*) (Figure 5-1). These factors correspond with the results of some recent researches in SSA and Ethiopia (see chapter 2). However, it conflicts with the findings from Ibrahim and Arulogun (2019) who do not find an effect of ethnicity in Nigeria. The factor *current number of children* can be categorized as another indirect determinant under the economic and investment theory since past reproductive experience influences the future investment parents are willing to make which in turn influences the FSP as well as the direct factors in Bongaarts' model (1978).

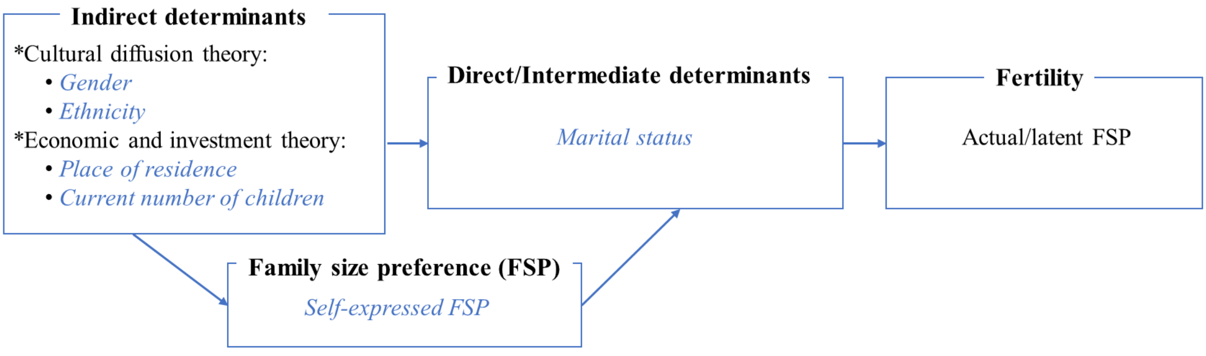


Figure 5-1: Factors influencing the heterogeneity in fertility preferences in this study according to Bongaarts' fertility model (1978)

Quality attribute

The analysis concerning the second research question doubting the existence of the quantity-quality trade-off suggests two elements. First, evidence from the **optimal mixtures** shows that for men, women and all the five LCs, the utility was maximized if most of the schooling mixture consisted of tertiary schooling ranging from 17% to 44% for the MXLM and LCM (Table 4-10 and 4-14). However, the preferred number of children finishing tertiary schooling is lower than self-expressed during the follow-up questions ranging from 89% to 100% for the MXLM and LCM (Table 4-6 and 4-11). This implies the existence of a trade-off. The main reasons given by the respondents for their preferred number of children during the follow-up questions were *economic*, followed by a *preference for quality of child-raising* and to *secure help in the household*. This is consistent with the findings from Teklu (2013) in Ethiopia. So, a lot of respondents themselves stated to have a high preference for quality of child-raising.

Second, the analysis of the **marginal rate of substitution** reveals that there is indeed a quantity-quality trade-off for some of the schooling attributes (Figure 4-5 and 8-13). However, for most attributes this only holds up to a certain level. The trade-off can be interpret in two ways: either it indicates that respondents are willing to have less children if more of their children can attain a higher educational level, or it indicates the high preference for number of children until respondents reach a certain ideal FSP. Since, in general, the ideal FSP lies around the endpoint of the trade-off, the second interpretation is more realistic. After this point, respondents are only willing to have more children, if more of them can finish a high educational level, indicating that the preference for quality predominates. To our knowledge, no research before has found the same result. Overall, Galor’s theoretical model (2012) suggests that family planning efforts focussing on increased availability, accessibility and affordability to quality can stimulate declines in quantity which is also found by Ito and Tanaka (2017). This can be accelerated by focussing on a decreasing gender wage gap as shown by Galor’s theoretical model (2012).

Gender attribute

The analysis related to the third research question (i.e. if fertility preferences and the quantity quality trade-off therein are gendered) reveals both a gender bias in the preference for the number of children and investment in education.

Concerning the **gender bias in preference for the number of children**, our results suggest four elements. First, the gendered sex ratio for the districts Chencha and Konso in the national data (Table 3-1) seems not to hold for our sample. Second, the self-expressed preferences reveal an overall higher preference for sons, consistent with the findings from Fuse (2010) (Table 4-6). Third, the DCE reveals that there is an overall higher preference for sons by men (Table 4-10 and 4-14). Women do not have a sex preference. Fourth, no heterogeneity in reasons for the self-expressed number of children is found except for girls (Table 4-8 and 4-12). Surprisingly, the second most frequent self-declared reason for a preference for girls is the continuation of the family lineage. However, during the focus group discussions, this reason was mainly quoted for a preference for boys. When girls marry, it is a cultural habit that they move to the family of the husband. A possible explanation could be that there is a possibility to extend the family lineage in terms of giving that child the same religion and ethnicity as the mother or via an extension of the family via the in-laws. Another clarification could be a coding error by the enumerators due to the low number of respondents preferring girls over boys (44 out of 426 respondents).

Further, concerning the possible **gender bias in investment in education**, this seems to hold. This was observed in the follow-up questions were respondents stated more frequently to prefer that their sons finish tertiary schooling (96%) compared to girls (92%) (Table 4-6). Additionally, the optimal mixtures in the DCE also revealed a gender bias in education for men (Table 4-10 and 4-14). Our modified model of Galor (2012) and de la Croix and Perrin (2016) shows that a shrinking gender education gap can be reached if either the fertility or the price of teaching decreases.

<u>Shortcomings</u>

Shortcomings of this study include the impossibility to calculate the response rate since the respondents were contacted orally by the head of the agricultural offices in each village. No full records of all the contacted persons were available. Second, it was beyond the scope of this research to include all possible indirect fertility determinants. After discussion with some key informants, only the main potential factors were selected. Third, no standard deviations could be obtained for the optimal number of children with the JMP Pro 14 software (SAS Institute Inc., 2018). Fourth, the results of the LCM differed for distinct starting values for the algorithm in the Stata 16 Software (Gould, 2016). Efforts were done to select the best model from 50 starting values. Fifth, the selection of the best LCM did not result in a clear characterization of the different LCs, although this was the best model that could be selected according to the BIC criterion and interpretability of the model. Sixth, our sample differs on a few aspects from the national data cautioning the extrapolation of our results to other regions.

6. Conclusion and policy recommendations

Fertility decline in human history is a complex enigma. Our aim is to add to the strand of literature focussing on SSA's exceptional population dynamics by performing a case study in Ethiopia. In particular we focussed on three gaps in the current state of the literature: i) the determinants explaining the heterogeneity in fertility preferences, ii) the existence of a quantity-quality trade-off and iii) the possible gender biased fertility preferences and quantity-quality trade-off. This was done via two types of data: a DCE (with attributes related to quantity, quality and gender, analysed via a MXLM and LCM) and follow-up questions. 426 respondents, between the age of 18 and 25, were interviewed to study ex ante fertility preferences (i.e. before completing their reproductive lifetime). This was complemented with information from four focus group discussions.

Our main results reveal that quantity and gender of children, and quality of child-raising are important in fertility preferences. The choice experiment revealed that the latent preference for the ideal family size is on average 5.98 children for men 5.62 children for women. This is more than one child higher than the self-expressed preferred family size (4.35 children for men and 4.29 for women). This gap is explained by the higher family size preference per se and not by an unmet need for contraception because of three reasons. First, 95% of the respondents stated to have access to contraception. Second, literature points out that there are also high levels of unrealized fertility in SSA. Third, the decline in TFR in Ethiopia started already before the government intervened via family planning policies focussing on contraception during history. In addition, the latent preference for the ideal family size is on average higher than the regional TFR of 5.2 children. This reflects that national data is possibly biased due to the low birth registration of 2.7%. Moreover, there is a large heterogeneity in the preferred number of children. The main determinants explaining the heterogeneity in fertility preferences in our study area are, as classified according to Bongaarts (1978), the *self-expressed FSP*, the direct factor *marital status* and indirect factors related to the cultural diffusion theory (*gender* and *ethnicity*), and economic and investment theory (*place of residence and current number of children*). So both the cultural diffusion theory, and the economic and investment theory seem to work in cooperation.

Further, our empirical analysis confirms the existence of a quantity-quality trade-off for most educational levels. Surprisingly, in general, this often only holds until respondents have reached their preferred family size. This implies that respondents are probably willing to give up schooling in favour of having more children. After reaching the preferred family size, respondents are only willing to have more children if these children can attain a higher educational level reflecting that the preference for quality predominates.

Next, our results reveal the presence of a gender bias concerning preference for the number of children and investment in education. According to the DCE, in general, the MXLM found that men have a stated preference for boys whereas women do not. Additionally, the LCM revealed that 88% of the respondents have a stated preference for boys and 12% for girls. This was also confirmed by the follow-up questions where 28% had a self-expressed preference for boys and 10% for girls.

Implications for current family planning policies and interventions are threefold. First, the gap between the stated and the self-expressed FSPs is explained by a higher prevailing FSP per se and not an unmet need for family planning services (and hence family planning efforts). Consequently, to achieve a reduction in current fertility rates, mainly efforts to reduce the FSP are needed, next to the current focus on contraception. We suggest that the demand for contraceptives will increase afterwards. Family planning interventions at best focus to decrease the higher prevailing desired family size of men, people who are married, have children, are living in the Gamo Gofa Zone or Konso district and belong to the Konso ethnic group. This can be done for example through mass media, health centres, education, etc. Further, improvements in birth registration are needed to make national data more reliable. Second, the existence of the trade-off implies that is possible to stimulate declines in quantity by increasing the availability, accessibility and affordability of education. This can be accelerated by focussing on a decreasing gender wage gap as shown by the model of Galor (2012). Third, to reduce the gender inequality in education, mainly men need to be targeted. Further, our modified model of Galor (2012) and de la Croix and Perrin (2016) shows that a shrinking gender education gap can be reached if either the number of children or the price of teaching decreases.

Suggestions for the future are to perform a more thorough analysis of the determinants of fertility preferences. This research shows that our DCE method is a valuable approach for this topic. Moreover, this thesis elicits that the application of a mixture-amount model is useful to analyse the quantity-quality trade-off due to the constraint that the number of children attaining a certain educational level need to sum up to the total number of children. Additionally, the non-linearity in the quantity-quality trade-off in different regions or conditions needs to be further analysed. Moreover, efforts are needed to make standard software more robust to work with non-linear data since our analysis of the LCM differed substantially from one starting value to another.

7. Bibliography

Aitchison, J., & Bacon-Shone, J. (1984). Log contrast models for experiments with mixtures. *Biometrika*, *71*(2), 323–330. https://www.jstor.org/stable/2336249%0D

Aleksandrovs, L., Goos, P., Dens, N., & de Pelsmacker, P. (2015). Mixed-media modeling may help optimize campaign recognition and brand interest. *Journal of Advertising Research*, *55*(4), 443–457. https://doi.org/10.2501/JAR-2015-025

Alemayehu, T., Haider, J., & Habte, D. (2010). Determinants of adolescent fertility in Ethiopia. *Ethiopian Journal of Health Development*, *24*(1). https://doi.org/10.4314/ejhd.v24i1.62942

Ali, M. M., Cleland, J., & Shah, I. H. (2012). *Causes and consequences of contraceptive discontinuation: evidence from 60 demographic and health surveys*.

Alidou, S., & Verpoorten, M. (2019). Family size and schooling in sub-Saharan Africa: testing the quantity-quality trade-off. *Journal of Population Economics*, *32*(4), 1353–1399. https://doi.org/10.1007/s00148-019-00730-z

Alvergne, A., Lawson, D. W., Clarke, P. M. R., Gurmu, E., & Mace, R. (2013). Fertility, parental investment, and the early adoption of modern contraception in rural Ethiopia. *American Journal of Human Biology*, *25*(1), 107–115. https://doi.org/10.1002/ajhb.22348

Aragaw, K. A. (2015). Application of logistic regression in determining the factors influencing the use of modern contraceptive among married women in Ethiopia. *American Journal of Theoretical and Applied Statistics*, *4*(3), 156–162. https://doi.org/10.11648/j.ajtas.20150403.21

Ariho, P., & Kabagenyi, A. (2020). Age at first marriage, age at first sex, family size preferences, contraception and change in fertility among women in Uganda: analysis of the 2006-2016 period. *BMC Women's Health*, *20*(1), 8. https://doi.org/10.1186/s12905-020-0881-4

Atake, E.-H., & Gnakou Ali, P. (2019). Women's empowerment and fertility preferences in high fertility countries in Sub-Saharan Africa. *BMC Women's Health*, *19*(54). https://doi.org/10.1186/s12905-019-0747-9

Bank, T. W. (2019). *Ethiopia gender diagnostic report*.

Barbour, R., & Kitzinger, J. (1999). *Developing focus group research: Politics, theory and practice* (1st ed.). Sage Publications.

Baschieri, A., Cleland, J., Floyd, S., Dube, A., Msona, A., Molesworth, A., Glynn, J. R., & French, N. (2013). Reproductive preferences and contraceptive use: a comparison of monogamous and polygamous couples in Northern Malawi. *Journal of Biosocial Science*, *45*, 145–166. https://doi.org/10.1017/S0021932012000569

Basu, D., & De Jong, R. (2010). Son targeting fertility behavior: some consequences and determinants. *Demography*, *47*(2), 521–536. https://doi.org/10.4324/9780203939451-14

Becker, G. S. (1960). An economic analysis of fertility. In *Demographic and economic change in developed countries* (pp. 209–240). Columbia University Press.

Becker, G. S., Murphy, K. M., & Tamura, R. (1990). Human capital, fertility, and economic growth. *Journal of Political Economy*, *98*(5), 12–37. https://doi.org/10.1086/261723

Becker, G. S., & Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of Political Economy*, *84*(4), 143–162. https://doi.org/10.2307/1831106

Becker, N. G. (1968). Models for the response of a mixture. *Journal of the Royal Statistical Society. Series B (Methodological)*, *30*(2), 349–358. https://doi.org/10.1111/j.2517-6161.1968.tb00735.x

Becker, N. G., & Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy*, *81*, 279–288. https://doi.org/10.1086/260166

Becker, N. S. (1978). Models and designs for experiments with mixtures. *Australian Journal of Statistics*, *20*(3), 195–208. https://doi.org/10.1111/j.1467-842X.1978.tb01102.x

Becker, S. O., Cinnirella, F., & Woessmann, L. (2010). The trade-off between fertility and education: evidence from before the demographic transition. *Journal of Economic Growth*, *15*, 177–204. https://doi.org/10.1007/s10887-010-9054-x

Behrman, J. A. (2015). Does schooling affect women's desired fertility? Evidence from Malawi, Uganda, and Ethiopia. *Demography*, *52*(3), 787–809. https://doi.org/10.1007/s13524-015-0392-3

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *5*(1), 289–300. https://www.jstor.org/stable/2346101

Bérenger, V., & Verdier-Chouchane, A. (2016). Child labour and schooling in South Sudan and Sudan: Is there a gender preference? *African Development Review*, *28*(S2), 177–190. https://doi.org/10.1111/1467-8268.12200

Berlie, A., & Alamerew, Y. (2018). Determinants of fertility rate among reproductive age women (15-49) in Gonji-Kollela District of the Amhara National Regional State, Ethiopia. *Ethiopian Journal of Health Development*, *32*(3), 1–12.

Bethlehem, J. (2009). The nonresponse problem. In *Applied Survey Methods* (pp. 209–248). John Wiley & Sons, Inc. https://doi.org/10.1002/9780470494998

Bhalotra, S., & Clarke, D. (2016). *The twin instrument* (No. 10405). https://ssrn.com/abstract=2886776

Bhat, C. R. (1997). Endogenous segmentation mode choice model with an application to intercity travel. *Transportation Science*, *31*(1), 34–48. https://doi.org/10.1287/trsc.31.1.34

Bielli, C., Berhanu, G., Isaias, A., & Orasi, A. (2001). *Population growth and environment in Ethiopia*.

Bingenheimer, J. B., & Raudenbush, S. W. (2004). Statistical and substantive inferences in public health: Issues in the application of multilevel models. *Annual Review of Public Health*, *25*(1), 53–77. https://doi.org/10.1146/annurev.publhealth.25.050503.153925

Bloor, M., Frankland, J., Thomas, M., & Robson, K. (2001). *Focus groups in social research* (1st ed.). Sage Publications.

Bongaarts, J. (1978). A Framework for Analyzing the Proximate Determinants of Fertility. *Population and Development Review*, *4*(1), 105–132. https://doi.org/10.2307/1972149

Bongaarts, J. (2011). Can family planning programs reduce high desired family size in Sub-Saharan Africa? *International Perspectives on Sexual and Reproductive Health*, *37*(4), 209–216. https://doi.org/10.1363/3720911

Bongaarts, J. (2015). Modeling the fertility impact of the proximate determinants: Time for a tune-up. *Demographic Research*, *33*(19), 535–560. https://doi.org/10.4054/DemRes.2015.33.19

Bongaarts, J. (2017). The effect of contraception on fertility: Is sub-Saharan Africa different? *Demographic Research*, *37*(6), 129–146. https://doi.org/10.4054/DemRes.2017.37.6

Bongaarts, J., & Casterline, J. (2013). Fertility transition: Is sub-Saharan Africa different? *Population and Development Review*, *38*, 153–168. https://doi.org/10.1111/j.1728-4457.2013.00557.x

Bongaarts, J., & Potter, R. E. (1983). *Fertility, biology, and behavior: An analysis of the proximate determinants* (1st ed.). Academic Press.

Bougma, M., LeGrand, T. K., & Kobiané, J.-F. (2015). Fertility limitation and child schooling in Ouagadougou: Selective fertility or resource dilution? *Studies in Family Planning*, *46*(2), 177–199. https://doi.org/10.1111/j.1728-4465.2015.00023.x

Burger, R. P., Burger, R., & Rossouw, L. (2012). The fertility transition in South Africa: a retrospective panel data analysis. *Development Southern Africa*, *29*(5), 738–755. https://doi.org/10.1080/0376835X.2012.731779

Burks, J. J., Randolph, D. W., & Seida, J. A. (2019). Modeling and interpreting regressions with interactions. *Journal of Accounting Literature*, *42*, 61–79. https://doi.org/10.1016/j.acclit.2018.08.001

Buyinza, F., & Hisali, E. (2014). Microeffects of women’s education on contraceptive use and fertility: The case of Uganda. *Journal of International Development*, *26*(6), 763–778. https://doi.org/10.1002/jid.2915

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: methods and applications* (1st ed.). Cambridge University Press.

Campbell, D., Boeri, M., & Doherty, E. (2015). Learning, fatigue and preference formation in discrete choice experiments. *Journal of Economic Behavior and Organization*, *119*, 345–363. https://doi.org/https://doi.org/10.1016/j.jebo.2015.08.018

Campbell, D., Hutchinson, W. G., & Scarpa, R. (2008). Incorporating discontinuous preferences into the analysis of discrete choice experiments. *Environmental and Resource Economics*, *41*(3), 401–417. https://doi.org/10.1007/s10640-008-9198-8

Caputo, V., Van Loo, E. J., Scarpa, R., Nayga, R. M., & Verbeke, W. (2018). Comparing serial, and choice task stated and inferred attribute non-attendance methods in food choice experiments. *Journal of Agricultural Economics*, *69*(1), 35–57. https://doi.org/10.1111/1477-9552.12246

Casterline, J. B., & Han, S. (2017). Unrealized fertility: Fertility desires at the end of the reproductive career. *Demographic Research*, *36*(14), 427–454. https://doi.org/10.4054/DemRes.2017.36.14

Central Statistical Agency. (2007a). *The 2007 population and housing census of Ethiopia: Statistical report for the Oromiya; Part I: Population size and characteristics*.

Central Statistical Agency. (2007b). *The 2007 population and housing census of Ethiopia: Statistical report for the SNNPR; Part I: Population size and characteristics*.

Central Statistical Agency. (2012). *Census 2007 Report*.

Central Statistical Agency, & ICF International. (2011). *Ethiopia demographic and health survey 2011*.

Central Statistical Agency, & ICF International. (2014). *Ethiopia mini demographic and health survey 2014*.

Central Statistical Agency, & ICF International. (2016). *Ethiopia demographic and health survey 2016*.

Central Statistical Agency, & ICF International. (2019). *Ethiopia mini demographic and health survey 2019*.

Channon, M. D., & Harper, S. (2019). Educational differentials in the realisation of fertility intentions: Is sub-Saharan Africa different? *PLoS ONE*, *14*(7). https://doi.org/10.1371/journal.pone.0219736

Chao, F., Gerland, P., Cook, A. R., & Alkema, L. (2019). Systematic assessment of the sex ratio at birth for all countries and estimation of national imbalances and regional reference levels. *Proceedings of The Royal Society B - Biological Sciences*, *116*(19), 9303–9311. https://doi.org/10.1073/pnas.1812593116

Chemhaka, G. B., & Odimegwu, C. O. (2019). The proximate determinants of fertility in Eswatini. *African Journal of Reproductive Health*, *23*(2), 65–75. https://doi.org/10.29063/ajrh2019/v23i2.7

Cinnirella, F. (2019). Marital fertility and investment in children's education. In *Cliometrics of the Family. Studies in Economic History* (pp. 33–54). Springer. https://doi.org/10.1007/978-3-319-99480-2_3

Cinnirella, F., & Streb, J. (2017). The role of human capital and innovation in economic development: evidence from post-Malthusian Prussia. *Journal of Economic Growth*, *22*(2), 193–227. https://doi.org/10.1007/s10887-017-9141-3

Clarke, D. (2018). Children and their parents: a review of fertility and causality. *Journal of Economic Surveys*, *32*(2), 518–540. https://doi.org/10.1111/joes.12202

Cleland, J., Bernstein, S., Ezeh, A., Faundes, A., Glasier, A., & Innis, J. (2006). Family planning: the unfinished agenda. *The Lancet*, *368*, 1810–1837. https://doi.org/10.1016/S0140

Coale, A. (1973). The Demographic Transition Reconsidered. *Proceedings of the International Population Conference*, 53–72.

Colleran, H., & Snopkowski, K. (2018). Variation in wealth and educational drivers of fertility decline across 45 countries. *Population Ecology*, *60*, 155–169. https://doi.org/10.1007/s10144-018-0626-5

Collins, A. T., Rose, J. M., & Hess, S. (2012). Interactive stated choice surveys: a study of air travel behaviour. *Transportation*, *39*(1), 55–79. https://doi.org/10.1007/s11116-011-9327-z

Cornell, J. A. (1971). Process variables in the mixture problem for categorized components. *Journal of the American Statistical Association*, *66*(333), 42–48. https://doi.org/10.1080/01621459.1971.10482215

Cornell, John A. (1981). *Experiments with mixtures : designs, models, and the analysis of mixture data*. Wiley-Blackwell.

Couper, M. P., Tourangeau, R., & Kenyon, K. (2004). Picture this: Exploring vsual effects in web surveys. *Public Opinion Quarterly*, *68*(2), 255–266. https://doi.org/10.1093/poq/nfh013

Davis, K., & Blake, J. (1956). Social structure and fertility: An analytic framework. *Economic*

Development and Cultural Change, *4*(3), 211–235. https://doi.org/10.1086/4497 14

de la Croix, D., & Perrin, F. (2016). *French fertility and education transition: Rational choice vs. cultural diffusion* (No. 2016–7).

de Leeuw, E. D., How, J. J., & Dillman, D. A. (2008). *International Handbook of Survey Methodology* (1st ed.). R. Taylor & Francis Group

Defo, B. K. (2011). The importance for the MDG4 and MDG5 of addressing reproductive health issues during the second secade of life: Review and analysis from times series data of 51 African countries. *African Journal of Reproductive Health*, *15*(2), 9–30. www.jstor.org/stable/45120802

Derringer, G., & Suich, R. (1980). Simultaneous optimization of several response variables. *Journal of Quality Technology*, *12*(4), 214–219. https://doi.org/10.1080/00224065.1980.11980968

Desai, J., & Tarozzi, A. (2011). Microcredit, family planning programs, and contraceptive behavior: Evidence from a field experiment in Ethiopia. *Demography*, *48*(2), 749–782. https://doi.org/10.1007/s13524-011-0029-0

Deshazo, J. R., & Fermo, G. (2002). Designing choice sets for stated preference methods: The effects of complexity on choice consistence. *Journal of Environmental Economics and Management*, *44*, 123–143. https://doi.org/doi:10.1006/jeem.2001.1199

Diebolt, C., & Perrin, F. (2019). A cliometric model of unified growth: Family organization and economic growth in the long run of history. In *Studies in Economic History* (pp. 7–31). Springer. https://doi.org/10.1007/978-3-319-99480-2_2

Draper, N. R., & St John, R. C. (1977). A mixtures model with inverse terms. In *Technometrics* (Vol. 19, Issue 1). https://www.jstor.org/stable/1268252

Easterlin, R. (1975). An economic framework for fertility analysis. *Studies in Family Planning*, *6*(3), 54–63.

Eastwood, R., & Lipton, M. (2011). Demographic transition in sub-Saharan Africa: How big will the economic dividend be? *Population Studies*, *65*(1), 9–35. https://doi.org/10.1080/00324728.2010.547946

Education Policy Data Center. (2019). *Ethiopia: Education overview*. https://www.epdc.org/country/ethiopia

Eissler, S., Thiede, B. C., & Strube, J. (2019). Climatic variability and changing reproductive goals in Sub-Saharan Africa. *Global Environmental Change*, *57*. https://doi.org/10.1016/j.gloenvcha.2019.03.011

Eliason, S., Baiden, F., Tuoyire, D. A., & Awusabo-Asare, K. (2018). Sex composition of living children in a matrilineal inheritance system and its association with pregnancy intendedness and postpartum family planning intentions in rural Ghana. *Reproductive Health*, *15*(1), 187. https://doi.org/10.1186/s12978-018-0616-2

Eloundou-Enyegue, P. M., & Giroux, S. C. (2012). Fertility transitions and schooling: From micro- to macro-level associations. *Demography*, *49*(4), 1407–1432. https://doi.org/10.1007/s13524-012-0131-y

Ethiopian Society of Population Studies. (2008). *Levels, trends and determinants of lifetime and desired fertility in Ethiopia: Findings from EDHS 2005*.

Farina, P., Gurmu, E., Hasen, A., & Maffioli, D. (2001). *Fertility and family change in Ethiopia*.

Federal Ministry of Education. (2015). *Education Sector Development Program V*.

Ferede, T. (2013). Multilevel modelling of modern contraceptive use among rural and urban population of Ethiopia. *American Journal of Mathematics and Statistics, 2013*(1), 1–16. https://doi.org/10.5923/j.ajms.20130301.01

Fern, E. F. (2001). *Advanced focus group research* (1st ed.). Sage Publications.

Fernihough, A. (2017). Human capital and the quantity–quality trade-off during the demographic transition. *Journal of Economic Growth*, *22*(1), 35–65. https://doi.org/10.1007/s10887-016-9138-3

Feyisa, A. D. (2019). *Household Survey on risk and resilience, soil and water conservation, land use and agricultural productivity in Southern Ethiopia*.

Fiebig, D. G., Keane, M. P., Louviere, J. J., & Wasi, N. (2010). The generalized multinomial logit model: Accounting for scale and coefficient heterogeneity. *Marketing Science*, *29*(3), 393–421. https://doi.org/10.1287/mksc.1090.0508

Flato, M. (2018). The differential mortality of undesired infants in Sub-Saharan Africa. *Demography*, *55*(1), 271–294. https://doi.org/10.1007/s13524-017-0638-3

Food and Agricultural Organization. (2018). *Ethiopia: Small family farms country factsheet*.

Fuse, K. (2010). Descriptive findings variations in attitudinal gender preferences for children across 50 less-developed countries. *Demographic Research*, *23*(36), 1031–1048. https://doi.org/10.4054/DemRes.2010.23.36

Galor, O. (2012). The demographic transition: Causes and consequences. *Cliometrica*, *6*(1), 1–28. https://doi.org/10.1007/s11698-011-0062-7

Galor, O., & Moav, O. (2002). Natural selection and the origin of rconomic growth. *Quarterly Journal of Economics*, *117*(4), 1133–1191. https://doi.org/10.1162/003355302320935007

Galor, O., & Weil, D. N. (1999). From Malthusian stagnation to modern growth. *The American Economic Review*, *89*(2), 150–154.

Galor, O., & Weil, D. N. (2000). Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. *The American Economic Review*, *90*(4), 806–828.

Gastwirth, J. L., Gel, Y. R., Hui, W. L. W., Lyubchich, V., Miao, W., & Noguchi, K. (2019). *lawstat: Tools for biostatistics, public policy, and law*. R Package Version 3.3. https://cran.r-project.org/package=lawstat

Gerland, P., Raftery, A. E., Ševčíková, H., Li, N., Gu, D., Spoorenberg, T., Alkema, L., Fosdick, B. K., Chunn, J., Lalic, N., Bay, G., Buettner, T., Heilig, G. K., & Wilmoth, J. (2014). World population stabilization unlikely this century. *Science*, *346*(6206), 234–237. https://doi.org/10.1126/science.1257469

Gibson, M. A., & Lawson, D. W. (2011). "Modernization" increases parental investment and sibling resource competition: evidence from a rural development initiative in Ethiopia. *Evolution and Human Behavior*, *32*(2), 97–105. https://doi.org/10.1016/j.evolhumbehav.2010.10.002

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of Grounded Theory* (1st ed.). Routledge. https://doi.org/10.4324/9780203793206

Goos, P., & Hamidouche, H. (2019). Choice models with mixtures: An application to a cocktail experiment. *Food Quality and Preference*, *77*, 135–146.

https://doi.org/10.1016/j.foodqual.2019.04.006

Goos, P., & Jones, B. (2011). *Optimal design of experiments: a case study approach* (1st ed.). John Wiley & Sons.

Goos, P., Jones, B., & Syafitri, U. (2016). I-Optimal Design of Mixture Experiments. *Journal of the American Statistical Association*, *111*(514), 899–911. https://doi.org/10.1080/01621459.2015.1136632

Gould, W. (2016). *Stata*. StataCorp.

Greenbaum, T. L. (1998). *The handbook for focus group research* (2nd ed.). Sage Publications.

Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. In *Transportation Research Part B: Methodological* (ITS-WP-02-08).

Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-Saliency theory of survey particiaption: description and an illustration. *The Public Opinion Quarterly*, *64*(3), 299–308. https://doi.org/10.2307/3078721

Gu, Y., Hole, A. R., & Knox, S. (2013). Fitting the generalized multinomial logit model in Stata. *The Stata Journal*, *13*(2), 382–397. https://doi.org/10.1177/1536867X1301300213

Günther, I., & Harttgen, K. (2016). Desired fertility and number of children born across time and space. *Demography*, *53*(1), 55–83. https://doi.org/10.1007/s13524-015-0451-9

Hall, R. W. (1999). Transportation Science. In *Handbook of transportation science* (1st ed., pp. 1–4). Springer US. https://doi.org/10.1007/978-1-4615-5203-1_1

Hayford, S. R., & Agadjanian, V. (2011). Uncertain future, non-numeric preferences, and the fertility transition: A case study of rural Mozambique. *Etude de La Population Africaine*, *25*(2), 419–439. https://doi.org/10.11564/25-2-239

Hensher, D. A. (2008). Joint estimation of process and outcome in choice experiments and implications for willingness to pay. *Journal of Transport Economics and Policy*, *42*(2), 297–322.

Hensher, D. A., & Greene, W. H. (2010). Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification. *Empirical Economics*, *39*(2), 413–426. https://doi.org/10.1007/s00181-009-0310-x

Hensher, D. A., Rose, J., & Greene, W. H. (2005a). The implications on willingness to pay of respondents ignoring specific attributes. *Transportation*, *32*(3), 203–222. https://doi.org/10.1007/s11116-004-7613-8

Hensher, D. A., Rose, J. M., & Greene, W. H. (2005b). *Applied choice analysis: a primer* (1st ed.). Cambridge University Press. https://doi.org/10.1007/9781316136232

Hess, S., & Hensher, D. A. (2010). Using conditioning on observed choices to retrieve individual-specific attribute processing strategies. *Transportation Research Part B: Methodological*, *44*, 781–790. https://doi.org/10.1016/j.trb.2009.12.001

Hess, S., & Rose, J. M. (2012). Can scale and coefficient heterogeneity be separated in random coefficients models? *Transportation*, *39*(6), 1225–1239. https://doi.org/10.1007/s11116-012-9394-9

Hess, S., & Stathopoulos, A. (2013). Linking response quality to survey engagement: a combined random scale and latent variable approach. *Journal of Choice Modelling*, *7*, 1–12. https://doi.org/10.1016/j.jocm.2013.03.005

Hess, S., Stathopoulos, A., Campbell, D., O'Neill, V., & Caussade, S. (2012). It's not that I don't care, I just don't care very much: Confounding between attribute non-attendance and taste heterogeneity. *Transportation, 40*(3), 583–607. https://doi.org/10.1007/s11116-012-9438-1

Hess, S., & Train, K. E. (2017). Correlation and scale in mixed logit models. *Journal of Choice Modelling*, *23*, 1–8. https://doi.org/10.1016/j.jocm.2017.03.001

Hinde, A. (1998). *Demographic methods* (1st ed.). Taylor & Francis Ltd.

Hole, A. R. (2010). *A discrete choice model with endogenous attribute attendance* (No. 2010006; Sheffield Economic Research Paper Series).

Hole, A. R. (2011). Attribute non-attendance in patients' choice of general practitioner appointment. *International Choice Modelling Conference,* 20.

Hong Il Yoo, L. (2019, November 18). *LCLOGIT2: Stata module to estimate latent class conditional logit models*. Boston College Department of Economics. https://econpapers.repec.org/RePEc:boc:bocode:s458616

Hossein, I., Saqib, N. U., & Haq, M. M. (2018). Scale heterogeneity in discrete choice experiment: An application of generalized mixed logit model in air travel choice. *Economics Letters*, *172*, 85–88. https://doi.org/10.1016/J.ECONLET.2018.08.037

Hoyos, D. (2010). The state of the art of environmental valuation with discrete choice experiments. *Ecological Economics*, *69*(8), 1595–1603. https://doi.org/10.1016/j.ecolecon.2010.04.011

Huber, J., & Train, K. E. (2001). On the similarity of classical and Bayesian estimates of individual mean partworths. *Marketing Letters*, *12*(3), 259–269.

Huber, Joel, & Zwerina, K. (1996). The importance of utility balance in efficient choice designs. *Journal of Marketing Research*, *33*(3), 307. https://doi.org/10.2307/3152127

Ibrahim, F. M., & Arulogun, O. S. (2019). Posterity and population growth: fertility intention among a cohort of Nigerian adolescents. *Journal of Population Research, 37*, 25–52. https://doi.org/10.1007/s12546-019-09230-z

Ito, T., & Tanaka, S. (2017). Abolishing user fees, fertility choice, and educational attainment. *Journal of Development Economics*, *130*, 33–44. https://doi.org/10.1016/j.jdeveco.2017.09.006

Kaplan, H. (1994). Evolutionary and wealth flows theories of fertility: Empirical tests and new models. *Population and Development Review*, *20*, 753–791. https://doi.org/10.2307/2137661

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/https://doi.org/10.2307/2291091

Kazeem, A., Jensen, L., & Stokes, C. S. (2010). School attendance in nigeria: Understanding the impact and intersection of gender, urban-rural residence, and socioeconomic status. *Comparative Education Review*, *54*(2), 295–319. https://doi.org/10.1086/652139

KC, S., & Lutz, W. (2014). Demographic scenarios by age, sex and education corresponding to the SSP narratives. *Population and Environment*, *35*(3), 243–260. https://doi.org/10.1007/s11111-014-0205-4

Kehlbacher, A., Balcombe, K., & Bennett, R. (2013). Stated attribute non-attendance in successive choice experiments. *Journal of Agricultural Economics*, *64*(3), 693–706. https://doi.org/10.1111/1477-9552.12021

Kessels, R., Jones, B., & Goos, P. (2019). Using Firth's method for model estimation and market segmentation based on choice data. *Journal of Choice Modelling, 31,* 1–21. https://doi.org/10.1016/j.jocm.2018.12.002

Kessels, R., Jones, B., Goos, P., & Vandebroek, M. (2008). *Recommendations on the use of Bayesian optimal designs for choice experiments* (KBI 0617).

Khademi, E., & Timmermans, H. (2012). Application of Mixture-Amount Choice Experiment for Accumulated Transport Charges. *Procedia - Social and Behavioral Sciences*, *54*, 483–492. https://doi.org/10.1016/j.sbspro.2012.09.766

Khuri, A. I. (2006). *Response surface methodology and related topics*. World Scientific Publishing Co. https://doi.org/10.1142/5915

Kjaer, T. (2005). *A review of the discrete choice experiment - with emphasis on its application in health care* (No. 1). Health Economic Papers.

Kodzi, I. A., Johnson, D. R., & Casterline, J. B. (2012). To have or not to have another child: Life cycle, health and cost considerations of Ghanaian women. *Social Science and Medicine*, *74*(7), 966–972. https://doi.org/10.1016/j.socscimed.2011.12.035

Kowalski, S. M., Cornell, J. A., & Vining, G. G. (2000). A new model and class of designs for mixture experiments with process variables. *Communications in Statistics - Theory and Methods, 29*(9–10), 2255–2280. https://doi.org/10.1080/03610920008832606

Kravdal, Ø., Kodzi, I., & Sigle-Rushton, W. (2013). Effects of the number and age of siblings on dducational transitions in Sub-Saharan Africa. *Studies in Family Planning*, *44*(3), 275–297. https://doi.org/10.1111/j.1728-4465.2013.00358.x

Krueger, R. A. (1988). *Focus groups: a practical guide for applied research* (5th ed.). Sage Publications.

Kuépié, M., Shapiro, D., & Tenikue, M. (2015). Access to schooling and staying in school in selected Sub-Saharan African countries. *African Development Review*, *27*(4), 403–414. https://doi.org/10.1111/1467-8268.12156

Lack, D. (1954). *The Natural Regulation of Animal Numbers* (1st ed.). Oxford University Press.

Laelago, T., Habtu, Y., & Yohannes, S. (2019). Proximate determinants of fertility in Ethiopia: An application of revised Bongaarts model. *Reproductive Health*, *16*(1), 13. https://doi.org/10.1186/s12978-019-0677-x

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Policital Economy, 74*(2), 132–157. https://www.jstor.org/stable/1828835

Lancsar, E., Fiebig, D. G., & Hole, A. R. (2017). Discrete choice experiments: a guide to model specification, estimation and software. *PharmacoEconomics*, *35*(7), 697–716. https://doi.org/10.1007/s40273-017-0506-4

Lancsar, E., Louviere, J., & Flynn, T. (2007). Several methods to investigate relative attribute impact in stated preference experiments. *Social Science and Medicine*, *64*(8), 1738–1753. https://doi.org/10.1016/j.socscimed.2006.12.007

Landry, A. (1987). Adolphe Landry on the demographic revolution. *Population and Development Review*, *13*(4), 731–740. https://www.jstor.org/stable/1973031

Lawson, D. W., & Borgerhoff Mulder, M. (2016). The offspring quantity-quality trade-off and human fertility variation. *Philosophical Transactions Series B - Biological Sciences, 371*(1692), 20150145. https://doi.org/10.1098/rstb.2015.0145

Leroy-Beaulieu. (1896). *Traité Théorique et Pratique d'Économie Politique* (1st ed.). Librairie Guillaumin et Cie. Lesthaeghe.

Livi Bacci, M. (2017). *A Concise History of World Population* (6th ed.). Wiley Blackwell.

Loomis, J. B. (2014). Strategies for overcoming hypothetical bias in stated preference surveys. *Journal of Agricultural and Resource Economics*, *39*(1), 34–46.

Louviere, J. J., Flynn, T. N., & Carson, R. T. (2010). Discrete Choice Experiments are not conjoint analysis. *Journal of Choice Modelling*, *3*(3), 57–72. https://doi.org/10.1016/S1755-5345(13)70014-9

Louviere, J. J., Street, D., Burgess, L., Wasi, N., & Islam, T. (2008). Modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information. *Journal of Choice Modelling*, *1*(1), 128–164. https://doi.org/10.1016/S1755-5345(13)70025-3

Louviere, J. J., Street, D., Carson, R., Ainslie, A., Deshazo, J. R., Cameron, T., Hensher, D., Kohn, R., & Marley, T. (2002). Dissecting the random component of utility: a conceptual framework for understanding the issues. *Marketing Letters*, *13*, 177–193. https://doi.org/10.1023/A:1020258402210

Mangham, L. J., Hanson, K., & Mcpake, B. (2009). How to do (or not to do): Designing a discrete choice experiment for application in a low-income country. *Health Policy and Planning*, *24*, 151–158. https://doi.org/10.1093/heapol/czn047

Mangiafico, S. (2020). *rcompanion: Functions to Support Extension Education Program Evaluation*. R Package Version 2.3.25. https://cran.r-project.org/package=rcompanion

Mani, S., Hoddinott, J., & Strauss, J. (2013). Determinants of schooling: Empirical evidence from rural Ethiopia. *Journal of African Economies*, *22*(5), 693–731. https://doi.org/10.1093/jae/ejt007

Manski, C. F. (1977). The structure of random utility models. *Theory and Decision*, *8*(3), 229–254. https://search.proquest.com/docview/1303217712?accountid=17215

Matthijs, K. (2012). *Bevolking : wie wat, waar, wanneer?* (2nd ed.). Acco.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (1st ed., pp. 105–142). Wiley.

McIntyre, D. M. (2014). Shared responsibilities for health: a coherent global framework for health financing. In *Final report of the Centre on Global Health Security Working group on Health Financing*. Chatham House.

Mekonnen, W., & Worku, A. (2011). Determinants of fertility in rural Ethiopia: the case of Butajira Demographic Surveillance System (DSS). *Public Health*, *11*(1), 782. https://doi.org/10.1186/1471-2458-11-782

Melstrom, R. T., Jayasekera, D. H., Boyer, T. A., & Jager, C. (2017). Scale heterogeneity in recreationists' decision making: Evidence from a site choice model of sport fishing. *Journal of Outdoor Recreation and Tourism, 18,* 81–87. https://doi.org/10.1016/j.jort.2017.02.007

Merton, R. K., & Kendall, P. L. (1946). The focused interview. *American Journal of Sociology*, *51*(6), 541–557. https://doi.org/10.1086/219886

Microsoft. (2010). *Microsoft Powerpoint* (No. 2010). Microsoft.

Minas, G. (2008). A review of the National Population Policy of Ethiopia. In *Digest of*

Ethiopia's National Policies, Strategies and Programs (pp. 23–45). Forum For Social Studies.

Ministry of Health [Ethiopia]. (2010). *Health and Health-Related Indicators*.

Ministry of Health [Ethiopia]. (2016a). *Costed Implementation Plan for Family Planning in Ethiopia 2015-2020*.

Ministry of Health [Ethiopia]. (2016b). *National Adolescent and Youth Health Strategy*.

Moav, O. (2005). Cheap children and the persistence of poverty. *The Economic Journal*, *115*, 88–110. https://doi.org/10.1111/j.1468-0297.2004.00961.x

Mogstad, M., & Wiswall, M. (2009). *How linear models can mask non-linear causal relationships: An application to family size and children's education* (No. 586).

Morgan, D. L. (1996). Focus groups. *Annual Review of Sociology*, *22*, 129–152. https://doi.org/10.1146/annurev.soc.22.1.129

Morgan, D. L. (1998). *Planning Focus Groups - Vol. 2* (1st ed.). Sage Publications.

Mortelmans, D. (2013). *Handboek kwalitatieve onderzoeksmethoden* (4th ed.). Acco.

Moya, C., Snopkowski, K., & Sear, R. (2016). What do men want? Re-examining whether men benefit from higher fertility than is optimal for women. *Philosophical Transactions Series B - Biological Sciences*, *371*. https://doi.org/10.1098/rstb.2015.0149

Mühlbacher, A. C., Zweifel, P., Kaczynski, A., & Johnson, F. R. (2015). Experimental measurement of preferences in health care using best-worst scaling (BWS): Theoretical and statistical issues. *Health Economics Review*, *6*(1), 1–12. https://doi.org/10.1186/s13561-015-0077-z

Ndagurwa, P., & Odimegwu, C. (2019). Decomposition of Zimbabwe's stalled fertility change: a two-sex approach to estimating education and employment effects. *Journal of Population Research*, *36*, 35–63. https://doi.org/10.1007/s12546-019-09219-8

Notestein, F. (1945). Population - The Long View. In T. W. Schultz (Ed.), *Food for the World* (1st ed., pp. 36–57). University Chicago Press.

Nylund-Gibson, K. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Structural Equation Modeling*, *14*(4), 535–569. https://doi.org/10.1080/10705510701575396

ODK Development Team. (2019). *Open Data Kit*.

Office of the Prime Minister [Ethiopia]. (1993). *National Population Policy of Ethiopia*.

Ogle, D. H., Wheeler, P., & Dinno, A. (2020). *FSA: Fisheries stock analysis*. R Package Version 0.8.30. https://github.com/droglenc/FSA

Olson, D. J., & Piller, A. (2013). Ethiopia: An emerging family planning success story. *Studies in Family Planning*, *44*(4), 445–459. https://doi.org/10.1111/j.1728-4465.2013.00369.x

Oumer, J. (2009). *The challenges of free primary education in Ethiopia*. University of Addis Ababa, International Institute for Educational Planning and UNESCO.

Pal, M., & Mandal, N. K. (2012). Optimum designs for estimation of parameters in a quadratic mixture-amount model. *Communications in Statistics-Theory and Methods*, *41*(4), 665–673. https://doi.org/10.1080/03610926.2010.529538

Perman, R., Ma, Y., Mcgilvray, J., & Common, M. (2011). *Natural Resource and*

Environmental Economics (4th ed.). Addison Wesley.

Piepel, G. F., & Cornell, J. A. (1985). Models for mixture experiments when the response depends on the total amount. *Technometrics*, *27*(3), 219. https://doi.org/10.2307/1269703

Population Reference Bureau. (2002). *Making the link: Population, health and the environment*. http://www.prb.org/Publications/Datasheets/2002/MakingTheLink.aspx

Powe, N. A., Garrod, G. D., & Mcmahon, P. L. (2005). Mixing methods within stated preference environmental valuation: choice experiments and post-questionnaire qualitative analysis. *Ecological Economics*, *52*(4). https://doi.org/10.1016/j.ecolecon.2004.06.022

Pradhan, U. K., Lal, K., Dash, S., & Singh, K. N. (2017). Design and analysis of mixture experiments with process variable. *Communications in Statistics - Theory and Methods*, *46*(1), 259–270. https://doi.org/10.1080/03610926.2014.990104

Prescott, P. (2004). Modelling in mixture experiments including interactions with process variables. *Quality Technology & Quantitative Management*, *1*(1), 87–103. https://doi.org/10.1080/16843703.2004.11673066

Pritchett, L. H. (1994). Desired fertility and the impact of population policies. *Population and Development Review*, *20*(1), 1–55. https://doi.org/10.2307/2137629

Proust, M. (2018). *JMP 14 Consumer Research*. SAS Institute Inc.

QSR International. (2018). *Nvivo* (12 Pro). QSR International.

R Core Team. (2020). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2002). Reliable estimation of generalized linear mixed models using adaptive quadrature. *Stata Journal*, *2*, 1–21. https://doi.org/10.1177/1536867X0200200101

Raghavarao, D., & Wiley, J. B. (2009). Conjoint measurement with constraints on attribute levels: a mixture-amount model approach. *International Statistical Review*, *77*(2), 167–178. https://doi.org/10.1111/j.1751-5823.2009.00077.x

Revelt, D., & Train, K. E. (1998). Mixed logit with repeated choices: Households' choices of appliance efficiency level. *Review of Economics and Statistics*, *80*(4), 647–657. https://doi.org/10.1162/003465398557735

Ringheim, K., Teller, C., & Sines, E. (2009). *Ethiopia at a crossroads: Demography, gender, and development*.

Roose, H., & Meuleman, B. (2014). *Methodologie van de sociale wetenschappen: een inleiding* (12th ed.). Academia Press.

Root, L., & Johnson-Hanks, J. (2016). Gender, honor, and aggregate fertility. *American Journal of Economics and Sociology*, *75*(4), 904–928. https://doi.org/10.1111/ajes.12159

Roser, M., Ritchie, H., & Ortiz-Ospina, E. (2019). *World Population Growth*. https://ourworldindata.org/world-population-growth

Rossi, P., & Rouanet, L. (2015). Gender Preferences in Africa: a Comparative Analysis of Fertility Choices. *World Development*, *72*, 326–345. https://doi.org/10.1016/j.worlddev.2015.03.010

Rowland, D. T. (2010). *Demographic methods and concepts* (2nd ed.). Oxford University Press.

Ruseckaite, A., Goos, P., & Fok, D. (2017). Bayesian D-optimal choice designs for mixtures. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *66*(2), 363–386. https://doi.org/10.1111/rssc.12174

Rutaremwa, G., Galande, J., Nviiri, H. L., Akiror, E., & Jhamba, T. (2015). The contribution of contraception, marriage and postpartum insusceptibility to fertility levels in Uganda: an application of the aggregate fertility model. *Fertility Research and Practice*, *1*(16). https://doi.org/10.1186/s40738-015-0009-y

Ruto, E., Garrod, G., & Scarpa, R. (2007). Valuing animal genetic resources: a choice modeling application to indigenous cattle in Kenya. *Agricultural Economics*, *38*(1), 89–98. https://doi.org/10.1111/j.1574-0862.2007.00284.x

Ryan, M., Watson, V., & Entwistle, V. (2009). Rationalising the 'irrational': a think aloud study of discrete choice experiment responses. *Health Economics*, *18*(3), 321–336. https://doi.org/10.1002/hec.1369

SAS Institue Inc. (2014). *SAS/QC 13.2 User's Guide: The OPTEX Procedure*.

SAS Institute Inc. (2013). *SAS 9.4* (9.4).

SAS Institute Inc. (2018). *JMP Pro 14* (14.0.0).

Scarpa, R., Gilbride, T. J., Campbell, D., & Hensher, D. A. (2009). Modelling attribute non-attendance in choice experiments for rural landscape valuation. *European Review of Agricultural Economics*, *36*(2), 151–174. https://doi.org/10.1093/erae/jbp012

Scarpa, R., Thiene, M., & Hensher, D. A. (2010). Monitoring choice task attribute attendance in nonmarket valuation of multiple park management services: does it matter? *Land Economics*, *86*, 817–839. https://doi.org/10.3368/le.86.4.817

Schreiber, J. B., & Pekarik, A. J. (2014). Technical note: Using latent class analysis versus K-means or hierarchical clustering to understand museum visitors. *Curator: The Museum Journal*, *57*(1), 45–59. https://doi.org/10.1111/cura.12050

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Shapiro, D., & Tenikue, M. (2017). Women's education, infant and child mortality, and fertility decline in urban and rural sub-Saharan Africa. *Demographic Research*, *37*(21), 670–699. https://doi.org/10.4054/DemRes.2017.37.21

Shenk, M. K., Towner, M. C., Kress, H. C., & Alam, N. (2013). A model comparison approach shows stronger support for economic models of fertility decline. *Proceedings of the National Academy of Sciences*, *110*(20), 8045–8050. https://doi.org/10.1073/pnas.1217029110

Snee, R. D. (1973). Techniques for the Analysis of Mixture Data. *Technometrics*, *15*(3), 517–528. https://doi.org/10.1080/00401706.1973.10489078

Street, D. J., Burgess, L., & Louviere, J. J. (2005). Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing*, *22*, 459–470. https://doi.org/10.1016/j.ijresmar.2005.09.003

Sugiura, N. (1978). Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics - Theory and Methods*, *7*(1), 13–26. https://doi.org/10.1080/03610927808827599

Swait, J. (1994). A structural equation model of latent segmentation and product choice for cross-sectional revealed preference choice data. *Journal of Retailing and Consumer*

Services, *1*(2), 77–89. https://doi.org/10.1016/0969-6989(94)90002-7

Taş, E. O., Reimão, M. E., & Orlando, M. B. (2014). Gender, ethnicity, and cumulative disadvantage in education outcomes. *World Development*, *64*, 538–553. https://doi.org/10.1016/j.worlddev.2014.06.036

Teklu, H., Sebhatu, A., & Gebreselassie, T. (2013). *Components of fertility change in Ethiopia: Further analysis of the 2000, 2005, and 2011 demographic and health surveys*.

Temel, T. (2013). Family size, human capital and growth: Structural path analysis of Rwanda. *Journal of Economic Development*, *38*(4), 39–73.

Tesfu, S. T., & Gurmu, S. (2013). Mother's Gender Preferences and Child Schooling in Ethiopia. *Atlantic Economic Journal*, *41*, 265–277. https://doi.org/10.1007/s11293-013-9366-2

The Ethiopian National Meteorological Services Agency. (2019). *Climatic and agroclimatic resources of Ethiopia*.

The World Bank. (2017). *World Population Prospects: The 2017 revision*.

The World Bank. (2020). *DataBank World Bank*. https://data.worldbank.org/indicator/

Thompson, W. S. (1929). Population. *American Journal of Sociology*, *34*, 959–975. https://doi.org/10.1086/214874

Thompson, W. S. (2003). The demographic transition model. In *Encyclopedia of Population* (1st ed., pp. 939–940). Macmillan Publishers.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, *101*(2), 266–270. https://doi.org/10.1037/h0070288

Timæus, I. M., & Moultrie, T. A. (2020). Pathways to low fertility: 50 years of limitation, curtailment, and postponement of childbearing. *Demography*, *57*, 267–296. https://doi.org/10.1007/s13524-019-00848-5

Towriss, C. A., & Timaeus, I. M. (2017). Schooling and intervals following a birth in Eastern Africa. *Population Studies*, *72*(1), 75–90. https://doi.org/10.1080/00324728.2017.1370121

Towriss, C. A., & Timaeus, I. M. (2018). Contraceptive use and lengthening birth intervals in rural and urban Eastern. *Demographic Research*, *38*, 2027–2052. https://doi.org/10.4054/DemRes.2018.38.64

Train, K. E. (2001). *A comparison of hierarchical bayes and maximum simulated likelihood for mixed logit*.

Train, K. E. (2008). EM Algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling*, *1*(1), 40–69. https://doi.org/10.1016/S1755-5345(13)70022-8

Trines, S. (2018). Education in Ethiopia. *SAGE Open*, *6*(1). https://doi.org/10.1177/2158244015624950

Trinitapoli, J., & Yeatman, S. (2018). The flexibility of fertility preferences in a context of uncertainty. *Population and Development Review*, *44*(1), 87–116. https://doi.org/10.1111/padr.12114

UNICEF. (2019). *Birth, Marriage and Death Registration in Ethiopia*. https://data.unicef.org/crvs/ethiopia/

United Nations. (2019a). *Human Development Reports*. http://hdr.undp.org/en/countries/profiles/ETH

United Nations. (2019b). *World Population Review*. http://worldpopulationreview.com/countries/ethiopia-population/

Upadhyay, U. D., & Karasek, D. (2012). Women's empowerment and ideal family size: an examination of DHS empowerment measures in Sub-Saharan Africa. *International Perspectives on Sexual and Reproductive Health*, *38*(2), 78–89. https://doi.org/10.1363/380

Van Der Pol, M., Currie, G., Kromm, S., & Ryan, M. (2014). Specification of the utility function in discrete choice experiments. *Value in Health*, *17*(2), 297–301. https://doi.org/10.1016/j.jval.2013.11.009

Van Hove, G., Bogdan, R. C., Biklen, S. K., Glesne, C., Neuman, W. L., & Howitt, D. (2014). *Qualitative Research for Educational Sciences* (1st ed.). Pearson Education Limited.

Van Lith, L. M., Yahner, M., & Bakamjian, L. (2013). Women's growing desire to limit births in sub-Saharan Africa: Meeting the challenge. *Global Health Science and Practice*, *1*(1), 97–107. https://doi.org/10.9745/GHSP-D-12-00036

Van Loo, E. J., Nayga, R. M., Campbell, D., Seo, H.-S., & Verbeke, W. (2018). Using eye tracking to account for attribute non-attendance in choice experiments. *European Review of Agricultural Economics*, *45*(3), 333–365. https://doi.org/10.1093/erae/jbx035

Vimefall, E., Andrén, D., & Levin, J. (2017). Ethnolinguistic background and enrollment in primary education: Evidence from Kenya. *African Development Review*, *29*(1), 81–91. https://doi.org/10.1111/1467-8268.12241

Vogl, T. S. (2016). Differential fertility, human capital, and development. *The Review of Economic Studies*, *83*(1), 365–401. https://doi.org/10.1093/restud/rdv026

Wado, Y. D., Gurmu, E., Tilahun, T., & Bangha, M. (2019). Contextual influences on the choice of long-acting reversible and permanent contraception in Ethiopia: a multilevel analysis. *PLoS ONE*, *14*(1). https://doi.org/10.1371/journal.pone.0209602

Werding, M., Editors Johannes Huinink, G., Ehrhardt, J., & Kohli, M. (2014). Children are costly, but raising them may pay: The economic approach to fertility. *Demographic Research*, *30*, 253–276. https://doi.org/10.4054/DemRes.2014.30.8

Willis, R. J. (1973). A new approach to the economic theory of fertility behavior. *Journal of Political Economy*, *81*(2), 14–64. https://doi.org/10.1086/260152

Wright, S. J., Vass, C. M., Sim, G., Burton, M., Fiebig, D. G., & Payne, K. (2018). Accounting for scale heterogeneity in healthcare-related discrete choice experiments when comparing stated preferences: a systematic review. *The Patient - Patient-Centered Outcomes Research*, *11*, 475–488. https://doi.org/10.1007/s40271-018-0304-x

Wurpts, I. C., & Geiser, C. (2014). Is adding more indicators to a latent class analysis beneficial or detrimental? Results of a Monte-Carlo study. *Frontiers in Psychology*, *5*(920). https://doi.org/10.3389/fpsyg.2014.00920

Yaya, S., Uthman, O. A., Ekholuenetale, M., & Bishwajit, G. (2018). Women empowerment as an enabling factor of contraceptive use in sub-Saharan Africa: a multilevel analysis of cross-sectional surveys of 32 countries. *Reproductive Health*, *15*(1), 214. https://doi.org/10.1186/s12978-018-0658-5

Zijlstra, T., Goos, P., & Verhetsel, A. (2019). A mixture-amount stated preference study on the mobility budget. *Transportation Research Part A - Policy*, *126*, 230–246. https://doi.org/10.1016/j.tra.2019.06.009

Znaniecki, F. (1934). *The method of sociology* (1st ed.). Farrar & Rinehart. https://doi.org/10.1177/000271623517700161

8. Appendix

Appendix 1: Drop-off, oral informed consent, respondents' characteristics and questions for the FGD

Drop-off

The results of these focus group discussions will be used for the following master thesis:

<u>Title of the research:</u> Fertility preferences and the trade-off between quantity and quality of child-raising: insights from southern Ethiopia

<u>Name and contact details of the researcher:</u>
Eva Boonaert
eva.boonaert@student.kuleuven.be
+32498060644

<u>Purpose and methodology of the research:</u>
The purpose is to understand the fertility preferences and the trade-off people make in the number of children and the quality of child-raising they can provide, including the study of gender differences. This is done via focus group discussions and a discrete choice experiment during the period July-October 2019 in the Gamo Gofa zone and the Segen People's zone of the Southern Nations, Nationalities, and People's Region in Ethiopia.

First name and name: ______________________

Age: _______ year

Residence place: __________________________

Educational level: _________________________

Profession: ______________________________

Do you currently have a partner? <u>yes/no</u>

If yes, name of the partner: ____________

Are you currently married? <u>yes/no</u>

Do you have children? <u>yes/no</u>

If yes, name of the children: ____

Figure 8-1: Drop-off before the focus group discussions

Oral Informed Consent before the focus group discussions

Welcome word

The purpose of these focus group discussions is to talk about family size planning and education around Arba Minch. This is Eva Boonaert, a master student from the KU Leuven University from Belgium who is conducting her thesis about fertility preferences and the trade-off people make in the number of children and the quality of child-raising they can provide, including the study of gender differences in southern Ethiopia. This is done via these focus group discussions and a discrete choice experiment during the period July-October 2019 in the Gamo Gofa zone and the Segen People's zone of the Southern Nations, Nationalities, and People's Region in Ethiopia. The results of the research can be used to improve current policies about child-raising.

During the group discussions, the purpose is to talk about the preferred family size and education in the region, because she wants to learn something about the cultural habits and traditions in the study area. The results of her research can be used to improve current policies about child-raising.

The group discussions will take maximum one and a half or two hours and anonymity of the data will be assured. You have the right to stop your participation in the study at any time. There are no false or correct answers. There will be a break in between with some drinks and snacks.

Are you willing to participate in these group discussions?

Can you understand the enumerator (same language)?

Figure 8-2: Oral informed consent before the focus group discussions

Table 8-1: Socio-demographic characteristics of the respondents of the focus group discussions. The educational levels refer to Figure 2-5.

Village	Gender	Pseudonym	Age	Educational level	Occupation	Partner	Married	Children
Morede	women	Abaynesh	22	10+3 (TVET)	Student	yes	yes	1
		Aster	22	9	Non-farm business	yes	yes	1
		Hibirnesh	18	5	Student	no	no	0
		Tunja	20	3	Housewife	no	no	1
		Kosora	19	10	Student	no	no	0
		Manalyshe	18	8	Student	no	no	0
		Kassawa	20	10+3 (TVET)	Student	no	no	0
	men	Bante	20	10	Farmer	yes	no	0
		Chencha	23	2	Farmer	yes	yes	2
		Gahano	24	10+3 (TVET)	Agricultural extension expert	yes	no	0
		Daniel	22	8	Farmer	yes	yes	1
		Mekaygno	20	7	Student	no	no	0
		Abinet	18	6	Student	no	no	no
Shelle Mela	women	Abebech	20	10	Housewife	yes	yes	0
		Adisse	25	9	Housewife	yes	yes	3
		Tarike	25	7	Housewife	yes	yes	2
		Burtukan	18	8	Student	no	no	0
		Marta	22	7	Housewife	yes	yes	2
		Amarech	25	6	Housewife	yes	yes	5
		Terafe	23	10	Housewife	yes	yes	1
	men	Bekele	22	10 + 3 (TVET)	Student	no	no	0
		Dima	22	12 + 1 (university)	Student	no	no	0
		Abera	20	12 + 3 (university) i.e. bachelor’s degree	Student	no	no	0
		Ayele	25	12 + 3 (university) i.e. bachelor’s degree	Teacher	yes	yes	0
		Zinabu	25	10	Farmer	yes	yes	2
		Anjulo	21	10	Preacher	yes	yes	2

Table 8-2: Questions asked during the focus group discussions

1. Brainstorming about your future family

"You will get all a piece of paper from me. I would like that you write or think about some short words that come into your mind when you think about your future family. Write them all down or think about them in your mind and afterwards, we will listen to everyone and give an overview on the board."

Additional questions:

Is family planning common in the region?

Why would people prefer a large number of children?

Why would people prefer a small number of children?

Why do people often prefer to have more boys?

2. Conceptual mapping about education

"Now I will ask you to write down all different types of education that exist here."

Additional questions:

At which level does most of the children quit school? Why?

3. Rating scales to start a discussion about the concepts used in the DCE

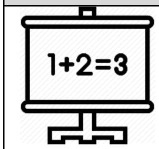
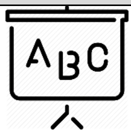
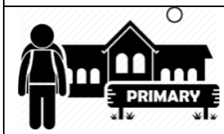
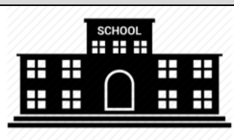
"Now, I will give you a list of different educational levels. Which level do you prefer for your own children?"

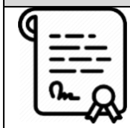
"Which is the minimum level that you want for your children?"

ORDER	TYPE OF SCHOOLING	LEVELS	CERTIFICATE
	primary school first cycle	grade 1-4	
	primary school second cycle	grade 5-8	Ethiopian School Leaving Certificate
	alternative basic education	grade 1-4	
	high school	grade 9-10	Ethiopian General Secondary Education Certificate
	preparatory school	grade 11-12	Ethiopian Higher Education Entrance Certificate
	technical and vocational schooling (TVET)	level 1-5	
	university	3-year program	
	university	5-year program	

4. Sorting pictures

"Now, I will give you some pictures concerning different levels of schooling. Indicate for each level which pictures the best resembles this."

Picture	Best picture = 1	Picture	Best picture = 1
No schooling			
			
			
Primary schooling			
			
			
			
Secondary schooling (high school and preparatory schooling)			
			
			
			

TVET or university			
			
			
			
			
Attaining a certificate for a certain educational level			
			
			
			

5. Forced choices

"Imagine, you will only get one child. Which educational level would you prefer for your child? Would you prefer to have a boy or a girl?"

"Imagine, you will get 12 children. Which educational levels would you prefer for your children? Does this differ for your boys or girls?"

6. Projective techniques for sensitive questions

"Now, I will give you a paper with some sentences. I will read them out loud so that you all understand the questions. I will ask you to complete the sentences or to think about your answers. Afterwards, we will go around and listen to everyone."

I **would/would not** [indicate] like to become married because________________________________

____

My ideal amount of children is____ of which ____ boys and ____ girls because________________

____

I **know/do not know** [indicate] something about the use of contraception.

→ If **yes**: I prefer to use____

7. Future thinking (if enough time)

"Assume that you would live in the future. How do you think a normal family will look like? How many children will they have? Which educational level will they obtain? You can write or paint this on your paper and then we will discuss this."

8. Debate (if enough time)

"Now we will split the group in 2. Each group has to defend an opinion. The left group prefers to have a few children, say 3 children. The right group prefers to have a lot of children, say 10. First you get 10 minutes to think with your group about possible arguments to defend your statement. Afterwards, you will have to convince the other group during 5 minutes and then, we will held a general discussion to see which group has convinced the others. For this discussion, your personal opinion does not count. You have to think from the view of the opinion I have given you."

Appendix 2: Experimental design of the DCE

Table 8-3: Experimental design with a the amount, x_{11}-x_{14} the number of girls attending no schooling, primary, secondary or tertiary schooling respectively and x_{21}-x_{24} the number of boys attending no schooling, primary, secondary or tertiary schooling respectively

		OPTION A											OPTION B										
BLOCK	SET	a	x_1	x_{11}	x_{12}	x_{13}	x_{14}	x_2	x_{21}	x_{22}	x_{23}	x_{24}	a	x_1	x_{11}	x_{12}	x_{13}	x_{14}	x_2	x_{21}	x_{22}	x_{23}	x_{24}
1	1	1	0	0	0	0	1	0	1	0	1	0	3	2	0	0	0	0	0	0	3	2	0
	2	1	1	0	1	0	0	0	0	0	1	1	1	0	1	0	3	0	1	2	1	0	1
	3	3	0	0	0	0	3	0	2	0	3	0	0	0	0	0	3	2	1	0	0	0	0
	4	3	0	0	0	0	3	0	0	2	3	0	6	0	0	6	4	0	0	0	6	0	0
	5	3	3	1	1	0	0	0	0	0	3	3	2	0	1	0	2	0	0	2	2	0	1
	6	2	0	0	0	0	2	0	1	1	2	0	6	0	0	6	2	0	0	0	6	0	0
	7	3	3	3	0	0	0	0	0	0	3	3	4	2	0	0	6	0	0	4	4	2	0
	8	5	3	0	2	0	2	0	1	0	5	3	8	0	4	0	4	0	0	4	8	0	4
	9	5	4	2	0	2	1	0	0	0	5	4	8	0	4	0	4	0	4	0	8	0	4
	10	5	4	0	2	2	1	0	1	0	5	4	4	2	0	0	1	1	0	0	4	2	0
	11	2	2	1	1	0	0	0	0	0	2	2	4	4	0	0	6	0	6	0	4	4	0
	12	9	6	3	0	0	3	0	0	3	9	6	6	0	0	4	6	0	0	4	6	0	0
2	1	5	1	0	0	0	4	3	0	0	5	1	4	0	0	3	1	0	0	0	4	0	0
	2	5	4	3	0	0	1	0	0	0	5	4	4	4	0	0	8	0	0	6	4	4	0
	3	1	0	0	0	0	1	0	0	1	1	0	3	3	0	0	1	0	0	1	3	3	0
	4	1	1	1	0	0	0	0	0	0	1	1	4	0	2	0	1	0	1	0	4	0	2
	5	10	6	0	4	0	4	0	0	4	10	6	0	0	0	0	3	1	1	0	0	0	0
	6	10	4	0	0	0	6	0	6	0	10	4	0	0	0	0	3	0	3	0	0	0	0
	7	8	2	0	0	2	6	2	4	0	8	2	8	0	8	0	4	0	4	0	8	0	8
	8	12	4	4	0	0	8	0	4	0	12	4	3	0	1	1	0	0	0	0	3	0	1
	9	3	2	0	0	1	1	0	0	1	3	2	4	0	0	0	8	4	0	4	4	0	0
	10	3	0	0	0	0	3	3	0	0	3	0	6	3	3	0	3	0	3	0	6	3	3
	11	3	1	0	1	0	2	0	1	0	3	1	1	0	0	0	4	0	3	0	1	0	0
	12	2	0	0	0	0	2	1	1	0	2	0	6	0	0	6	6	0	6	0	6	0	0
3	1	2	1	0	0	1	1	0	0	1	2	1	1	1	0	0	4	0	0	2	1	1	0
	2	3	3	0	2	0	0	0	0	0	3	3	2	0	1	1	0	0	0	0	2	0	1
	3	2	1	1	0	0	1	1	0	0	2	1	0	0	0	0	3	1	0	2	0	0	0
	4	3	0	0	0	0	3	0	1	1	3	0	2	2	0	0	3	0	1	0	2	2	0
	5	10	6	4	0	2	4	0	0	0	10	6	3	0	3	0	1	0	0	0	3	0	3
	6	5	1	0	0	0	4	0	0	3	5	1	4	4	0	0	8	0	0	4	4	4	0
	7	3	1	0	0	1	2	1	0	0	3	1	4	0	0	3	1	0	0	0	4	0	0
	8	3	0	0	0	0	3	2	0	0	3	0	3	1	2	0	2	0	0	0	3	1	2
	9	12	4	0	4	0	8	0	4	0	12	4	1	0	1	0	1	1	0	0	1	0	1
	10	5	1	0	1	0	4	2	0	0	5	1	4	0	0	0	8	0	4	4	4	0	0
	11	5	2	0	0	1	3	2	0	0	5	2	8	0	8	0	4	0	0	4	8	0	8
	12	2	1	1	0	0	1	0	0	1	2	1	6	0	0	6	3	0	0	3	6	0	0
4	1	10	4	4	0	0	6	0	4	0	10	4	2	0	0	0	3	0	0	3	2	0	0
	2	2	1	0	1	0	1	0	0	1	2	1	4	0	0	4	8	0	0	4	4	0	0
	3	1	1	0	0	1	0	0	0	0	1	1	4	0	0	4	8	8	0	0	4	0	0
	4	3	0	0	0	0	3	1	0	1	3	0	6	6	0	0	2	0	0	0	6	6	0
	5	10	6	0	0	2	4	4	0	0	10	6	1	0	1	0	2	0	0	1	1	0	1
	6	5	1	0	0	0	4	0	4	0	5	1	4	0	0	2	1	0	1	0	4	0	0

	7	10	6	0	0	2	4	0	0	4	10	6	8	0	0	6	4	4	0	0	8	0	0
	8	9	6	0	3	0	3	3	0	0	9	6	2	2	0	0	2	0	2	0	2	2	0
	9	8	2	2	0	0	6	4	0	2	8	2	4	0	0	4	8	0	4	0	4	0	0
	10	8	4	2	0	0	4	4	0	0	8	4	4	2	0	0	1	0	1	0	4	2	0
	11	3	1	1	0	0	2	1	0	0	3	1	4	0	4	0	8	4	0	0	4	0	4
	12	12	4	0	0	0	8	4	4	0	12	4	6	0	3	3	3	0	0	0	6	0	3
5	1	5	2	0	0	0	3	3	0	0	5	2	8	4	4	0	4	0	0	0	8	4	4
	2	12	8	0	4	4	4	0	0	0	12	8	1	0	0	1	4	0	2	0	1	0	0
	3	1	0	0	0	0	1	1	0	0	1	0	8	8	0	0	4	4	0	0	8	8	0
	4	5	3	3	0	0	2	0	0	0	5	3	2	1	0	1	0	0	0	0	2	1	0
	5	4	2	0	0	1	2	0	2	0	4	2	6	0	0	6	4	4	0	0	6	0	0
	6	10	6	4	0	0	4	0	4	0	10	6	3	1	0	1	0	0	0	0	3	1	0
	7	3	2	0	0	2	1	1	0	0	3	2	4	0	4	0	6	0	0	2	4	0	4
	8	3	3	0	0	2	0	0	0	0	3	3	4	0	2	0	1	0	0	1	4	0	2
	9	12	4	0	4	0	8	8	0	0	12	4	1	0	0	1	2	0	0	1	1	0	0
	10	8	4	0	4	0	4	0	4	0	8	4	8	0	4	0	4	4	0	0	8	0	4
	11	8	2	2	0	0	6	0	0	6	8	2	0	0	0	0	2	1	0	1	0	0	0
	12	5	1	0	0	0	4	0	3	0	5	1	4	0	3	0	1	0	0	0	4	0	3

Appendix 3: Follow-up questions

ANSWER SHEET

Contact persons:
Eva Boonaert: +251996695432
Ashenafi Duguma Feyisa: +251911540111

CONTACT DESCRIPTION

QUESTION	ENCIRCLE	
1. Is it possible to make contact with the sampled respondent?	Yes	→ Explain the purpose of the research + question 2
	No	→ Reason (see Code 1): ____________
2. What is the age of the respondent?	_____ year	→ Question 3
3. Is the respondent between the age of 18 and 25 (=eligible)?	Yes	→ Question 4
	No	→ END OF THE SURVEY
4. Is the respondent willing to participate in the survey?	Yes	→ Question 5
	No	→ Reason (see Code 2): ____________
5. Is there a common language to perform the survey?	Yes	→ Perform the survey
	No	→ A translator from the kebele is needed

Code 1

a) Address is not valid (unoccupied, demolished…)
b) Respondent is not at home
c) Respondent has moved

Code 2 [multiple answers possible]

a) Bad timing
b) Not interested
c) Too complicated for me
d) I do not trust surveys
e) Waste of time
f) I do not want to give personal information
g) Participated too often in surveys
h) Previous bad experience with surveys
i) Refusal because partner/family do not give permission to participate
j) Mental or physical problems
k) Other reason: explain

GENERAL INFORMATION

1 Name of the respondent: ____________________
2 District/Woreda: ____________________
3 Village/Kebele: ____________________
4 Telephone number: ____________________
5 Date: ____________________
6 Interviewer name: ____________________
7 Start time: _______

DEMOGRAPHIC CHARACTERISTICS

1 Sex: male/female
2 Level of education (see Code 3): __________
3 Years of education: ____________________
4 Principal occupation (see Code 4): ____________________
5 What is your religion? ______________
6 What is your residence history of the household head? __________
7 To which ethnic group do you belong to? ______________

Code 3

a) No education
b) No formal education but can read & write
c) Primary schooling (grade 1-8)
d) Secondary schooling: 1st cycle (grade 9-10)
e) Secondary schooling: 2nd cycle (grade 11-12)
f) TVET
g) University
h) Adult education
i) Religious education

Code 4

a) Student
b) Family laborer on the own farm
c) Farmer on own farm
d) Hired farm laborer
e) Family laborer in the household
f) Household wife in own household
g) Family laborer in a non-farm business
h) Own non-farm business
i) Hired servant
j) Civil servant
k) Entrepreneur
l) Looking for a job
m) Religious person
n) Charity worker
o) Unable to work
p) Other: specify

CHOICE EXPERIMENT

Type of cards		Number of the choice card	Answer	
Test cards → explain how to perform the choice experiment		Test card 1	A	B
		Test card 2	A	B
Block nr. _____	Choice experiment with the 12 cards → shuffle them		A	B
			A	B
			A	B
			A	B
			A	B
			A	B
			A	B
			A	B
			A	B
			A	B
			A	B
			A	B
	Duplicates (2 of the 12 cards → shuffle them)		A	B
			A	B

Figure 8-3: Example of the answer sheet (frontside). This was digitized in the tablets.

Which attribute did you ignored? **[multiple answers possible]**
☐ Number of children
☐ Gender
☐ No schooling
☐ Educational level
☐ All options are considered

Which attribute do you find the least important?
☐ Number of children
☐ Gender
☐ No schooling
☐ Educational level

ADDITIONAL QUESTIONS

1. Are you **married** (Code 5)?
   - If yes (=married): At what age? ______ year
     - Is this your first marriage? Yes/No
     - [for women only]: Are you the only women? Yes/No
       - If no: How many wives does your husband have? _______
       - How many wives did the household head have before you? ___
   - If **no** (=not married): Do you prefer to become married? Yes/No
     - If yes: At what age do you prefer to become married? ______ year

Code 5
a) Yes, married
b) No, single
c) No, divorced
d) No, widowed

2. Do you have **children**?: Yes/No
   - If yes (=children): How many children do you have (alive)? ____ of which ____ boys and ____ girls
     - At what age did you get your first child? _____ year
     - Do you have children who have deceased? Yes/No
       - If yes: Can you tell me how many? ____sons & ____daughters
     - What is your ideal number of children? ____ of which ____ boys and ____ girls
       - What influences your ideal number of children (see Code 6) **[multiple answers possible]**? _________
       - [If the respondent has a preference for more boys or girls]: Why do you prefer more boys (or girls) (see Code 7) **[multiple answers possible]**? _________
     - What is the ideal level of future education for your boys (see Code 8)? ________
     - What is the ideal level of future education for your girls (see Code 8)? ________
   - If **no** (=no children): What is your ideal number of children? ____ of which ____ boys and ____ girls
     - At what age do you prefer to have your first child? _____year
       - What influences your ideal number of children (see Code 6) **[multiple answers possible]**? _________
       - [If the respondent has a preference for more boys or girls]: Why do you prefer more boys (or girls) (see Code 7) **[multiple answers possible]**? _________
     - What is the ideal level of education for your future boys (see Code 8)? ________
     - What is the ideal level of education for your future girls (see Code 8)? ________

Code 6
a) To secure help in the household
b) To secure caretaking when parents become older
c) Wealth status
d) Decision of spouse
e) Gender preference (certain girl/boy ratio)
f) Expected from the community
g) Expected from my religion
h) Economic reason
i) Preference for quality of child-raising
j) Health of the mother
k) Other reason: explain

Code 7
a) To secure help in the household
b) Wealth status
c) Gender preference (certain girl/boy ratio)
d) Expected from the community
e) Continuation of the family lineage
f) Other reason: explain

Code 8
a) Primary schooling
b) Secondary schooling: 1st cycle
c) Secondary schooling: 2nd cycle
d) TVET
e) University
f) Does not matter

3. Do you have knowledge about **contraception**?: Yes/No
   - If yes (=knowledge): Which methods do you know (see Code 8) **[multiple answers possible]**? _________
     - Which method do you preferred the most (see Code 9)? ____________
     - Have you ever used contraception? Yes/No
       - If yes: Are you currently using contraception? Yes/No
     - Can you have access to contraception in this kebele? Yes/No
       - If yes: Is there sometimes a shortage? No/once in 2 or 3 months/once a month
       - If **no**: Would you prefer to have more access to it? Yes/No

Code 9
a) Depo (injection e.g. jadelle)
b) Contraceptive pills
c) Implants
d) Loop or intra-uterine device (IUD)
e) Condom
f) Tubal/ dicotomay (long-term method for women)
g) Vasectomy (long-term method for men)
h) Lactational amenorrhea method (LAM)
i) Post pil
j) Traditional/natural methods
k) Ohter: explain

4. End time: _______

AFTERWARDS: INTERVIEWER OBSERVATIONS

Remarks about the respondent: ____

Remarks about the choice experiment: ______________________________________

Remarks about the survey: ____

THANK YOU VERY MUCH FOR YOUR COOPERATION AND YOUR TIME!

Figure 8-4: Example of the answer sheet (backside). This was digitized in the tablets.

Appendix 4: Model selection for the DCE

*Table 8-4: Comparison of the AICc, BIC, -2*log-likelihood and -2*Firth log-likelihood for different models for the model selection. For some higher order powers, no convergence of the iterations was achieved if the number of parameters exceeded the maximum of 120 degrees of freedom of our blocked choice design (i.e. 5 blocks of 12 choice cards with each 2 alternatives).*

Scheffé model	**Amount effect**		**Number of parameters**	**Quality measures**		
				AIC_c	**BIC**	**-2*log -likelihood**
MXLM						
First order or linear	*Multiplicative*	*Linear*	15	460.66	656.54	400.30
		Quadratic	23	239.87	**539.92**	147.02
		3rd power	31	143.62	547.63	18.07
		4th power	39	169.79	677.57	11.35
		5th power	47	202.39	813.72	10.84
		6th power	55	234.81	949.48	9.93
		7th power	63	267.87	1085.67	9.47
Second order or quadratic	*Multiplicative*	*Linear*	70	169.20	625.12	27.23
		Quadratic	105	226.51	908.91	12.07
LCM 2 Classes						
First order or linear	*Multiplicative*	*Linear*	31	5243.10	5445.49	5180.71
		Quadratic	47	4774.96	5081.51	4680.07
		3rd power	63	4372.37	4782.88	4244.78
		4th power	79	4278.86	4793.11	4118.35
		5th power	95	3268.09	**3885.89**	3074.46
		6th power	111	/	/	/
		7th power	127	/	/	/
Second order or quadratic	*Multiplicative*	*Linear*	141			
		Quadratic	211			
LCM 3 Classes						
First order or linear	*Multiplicative*	*Linear*	47	4890.37	5196.92	4795.48
		Quadratic	71	4466.93	4929.34	4322.91
		3rd power	95	4281.44	**4899.23**	4087.81
		4th power	119	4575.72	5348.42	4332.01
		5th power	143	/	/	/
		6th power	167	/	/	/
		7th power	191	/	/	/
Second order or quadratic	*Multiplicative*	*Linear*	212	/	/	/
		Quadratic	317	/	/	/
LCM 4 Classes						
First order or linear	*Multiplicative*	*Linear*	63	4791.71	**5202.22**	4664.12
		Quadratic	95	4599.82	5217.62	4406.19
		3rd power	127	/	/	/
		4th power	159	/	/	/
		5th power	191	/	/	/
		6th power	223	/	/	/
		7th power	255	/	/	/
Second order or quadratic	*Multiplicative*	*Linear*	283	/	/	/
		Quadratic	423	/	/	/
LCM 5 Classes						
First order or linear	*Multiplicative*	*Linear*	79	4725.14	5239.39	4564.63
		Quadratic	119	4327.71	**5100.41**	4084.00
		3rd power	159	/	/	/
		4th power	199	/	/	/
		5th power	239	/	/	/
		6th power	279	/	/	/
		7th power	319	/	/	/
Second order or quadratic	*Multiplicative*	*Linear*	354	/	/	/
		Quadratic	529	/	/	/
LCM 6 Classes						
First order or linear	*Multiplicative*	*Linear*	95	4639.17	**5256.97**	4445.54
		Quadratic	143	/	/	/
		3rd power	191	/	/	/
		4th power	239	/	/	/
		5th power	287	/	/	/
		6th power	335	/	/	/
		7th power	383	/	/	/
Second order or quadratic	*Multiplicative*	*Linear*	425	/	/	/
		Quadratic	635	/	/	/

Note: The best option according to the BIC (which penalizes more for more complex models and is therefore recommended by Schwarz (1978)) is indicated in bold.

Appendix 5: Results of the LCM

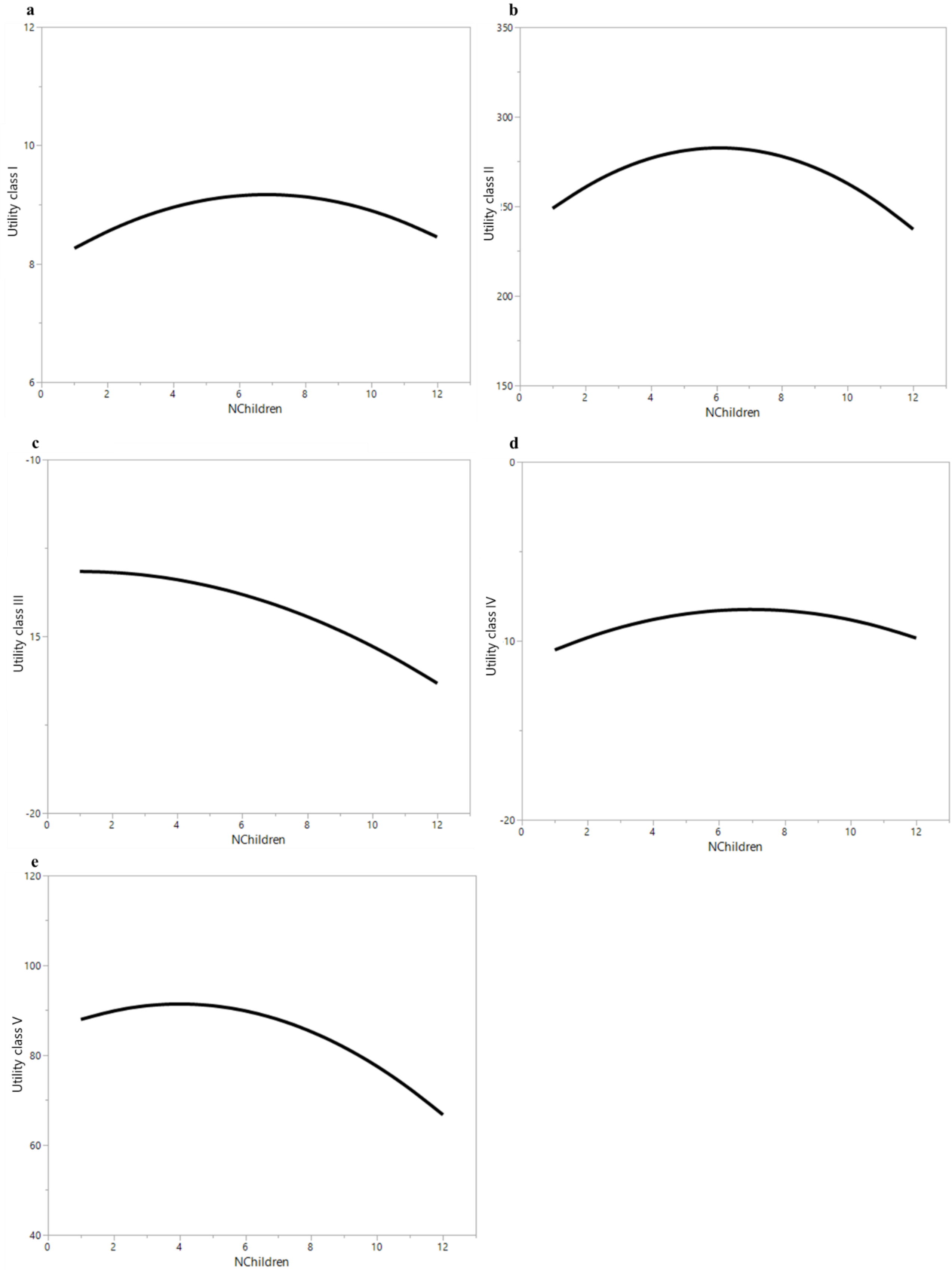


*Figure 8-5: The respondent's utility in function of the number of children for LC I to V corresponding with graphs **a** to **e** respectively*

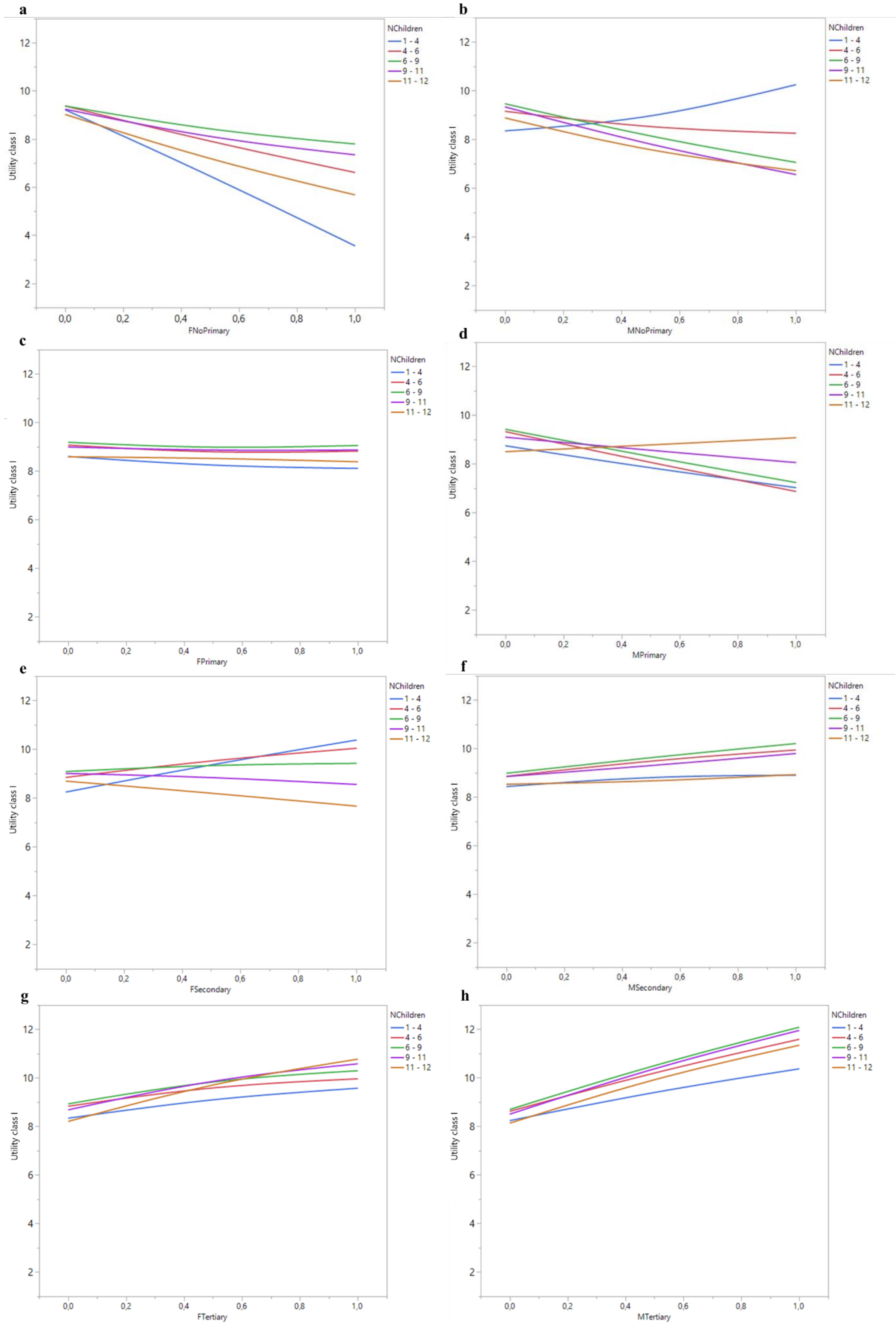


Figure 8-6: Visualization of ***LC I****'s utility in function of the different schooling components for different ranges of the number of children. Note: all results need to be interpreted with respect to FNoPrimary. The explanation of the variables can be found in Table 3-2.*

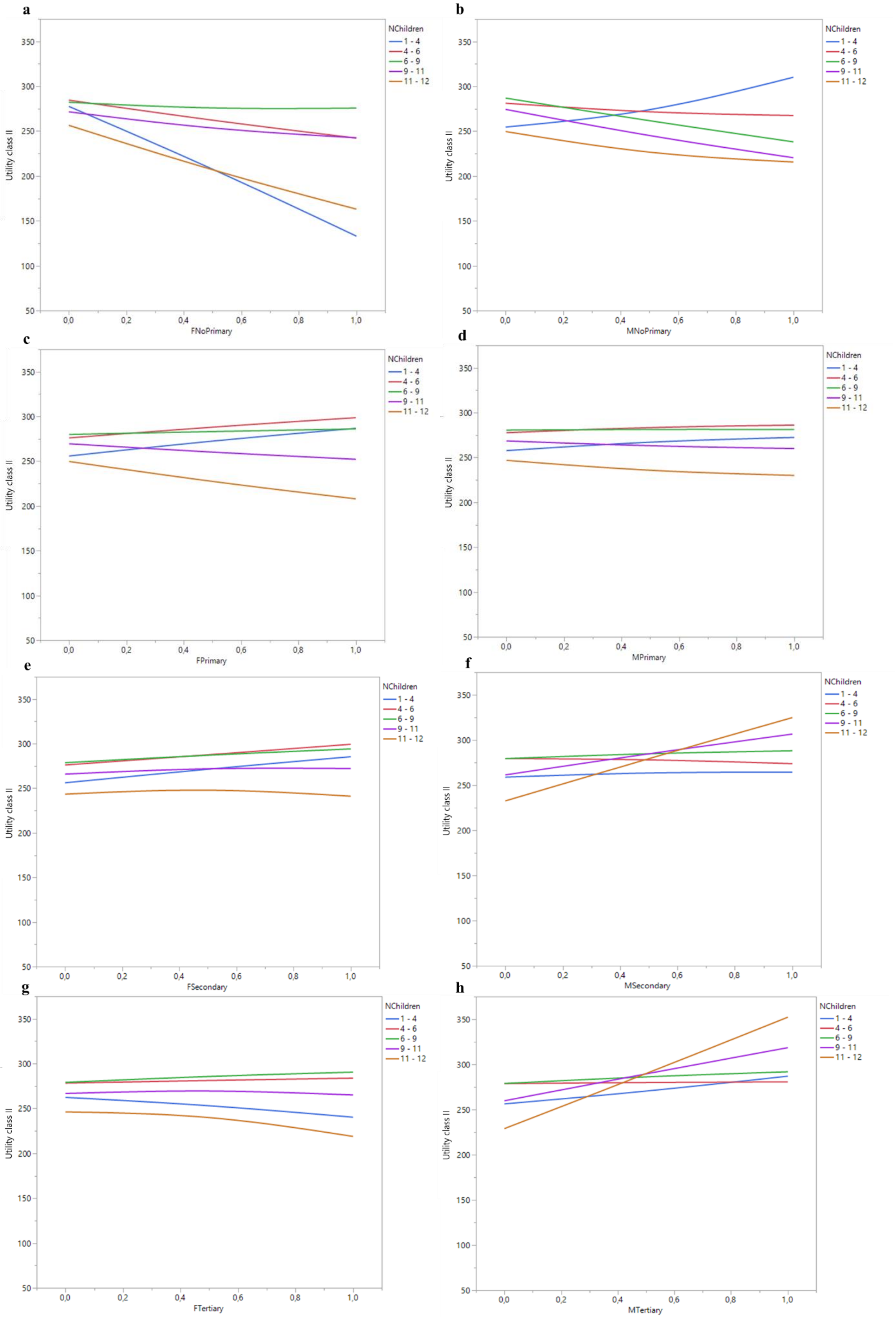


Figure 8-7: Visualization of ***LC II****'s utility in function of the different schooling components for different ranges of the number of children. Note: all results need to be interpreted with respect to FNoPrimary. The explanation of the variables can be found in Table 3-2.*

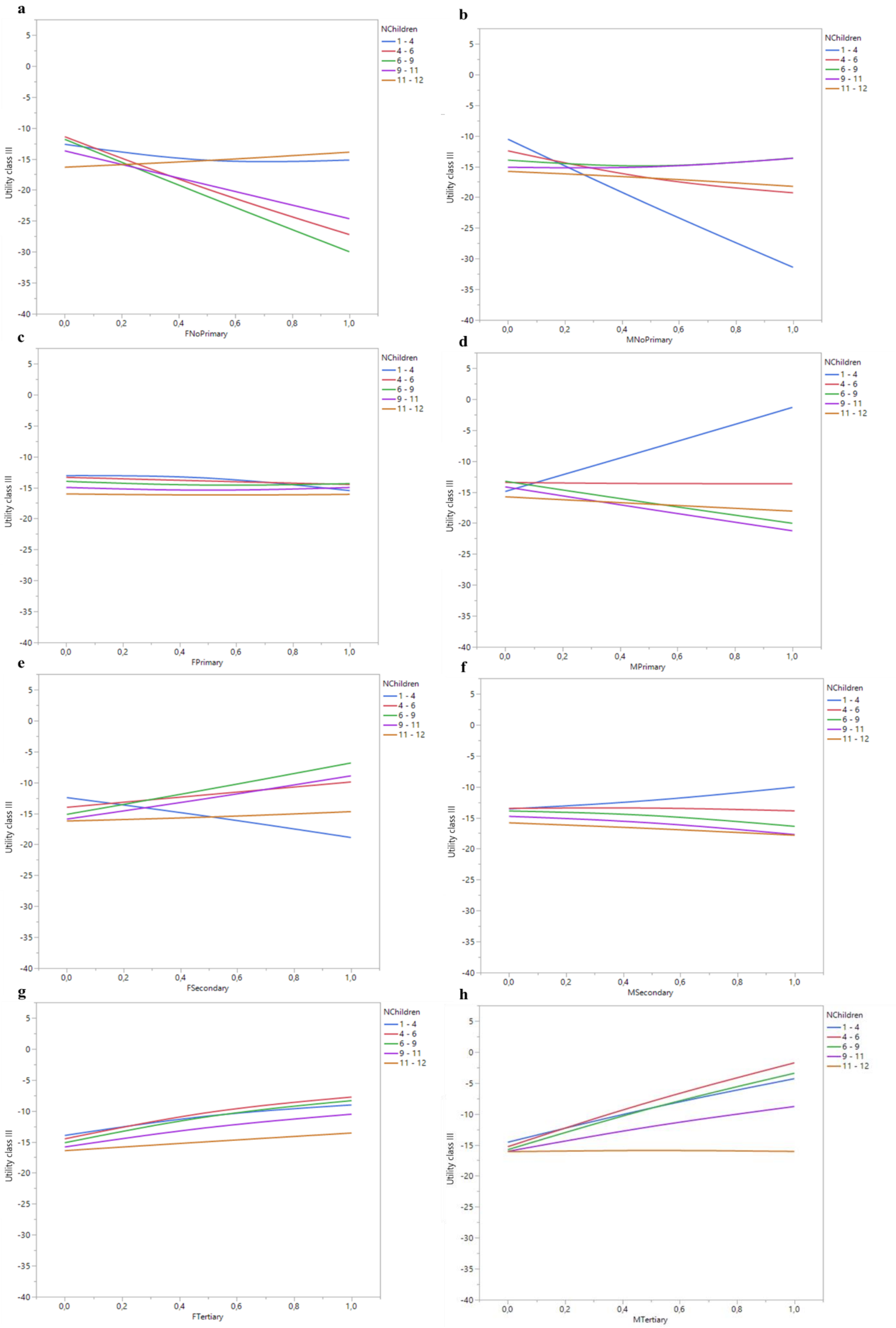


*Figure 8-8: Visualization of **LC III**'s utility in function of the different schooling components for different ranges of the number of children. Note: all results need to be interpreted with respect to FNoPrimary. The explanation of the variables can be found in Table 3-2.*

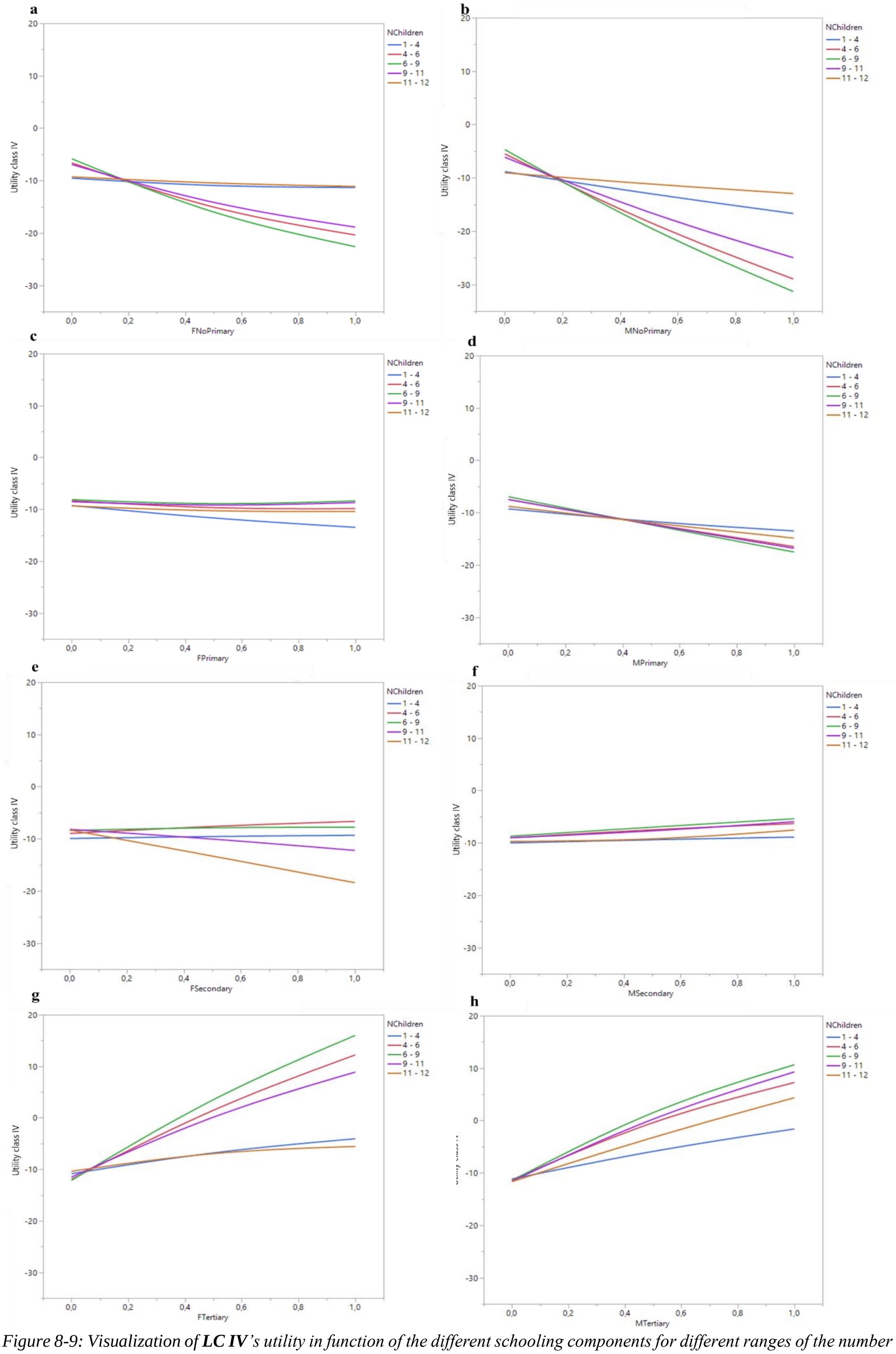


*Figure 8-9: Visualization of **LC IV**'s utility in function of the different schooling components for different ranges of the number of children. Note: all results need to be interpreted with respect to FNoPrimary. The explanation of the variables can be found in Table 3-2.*

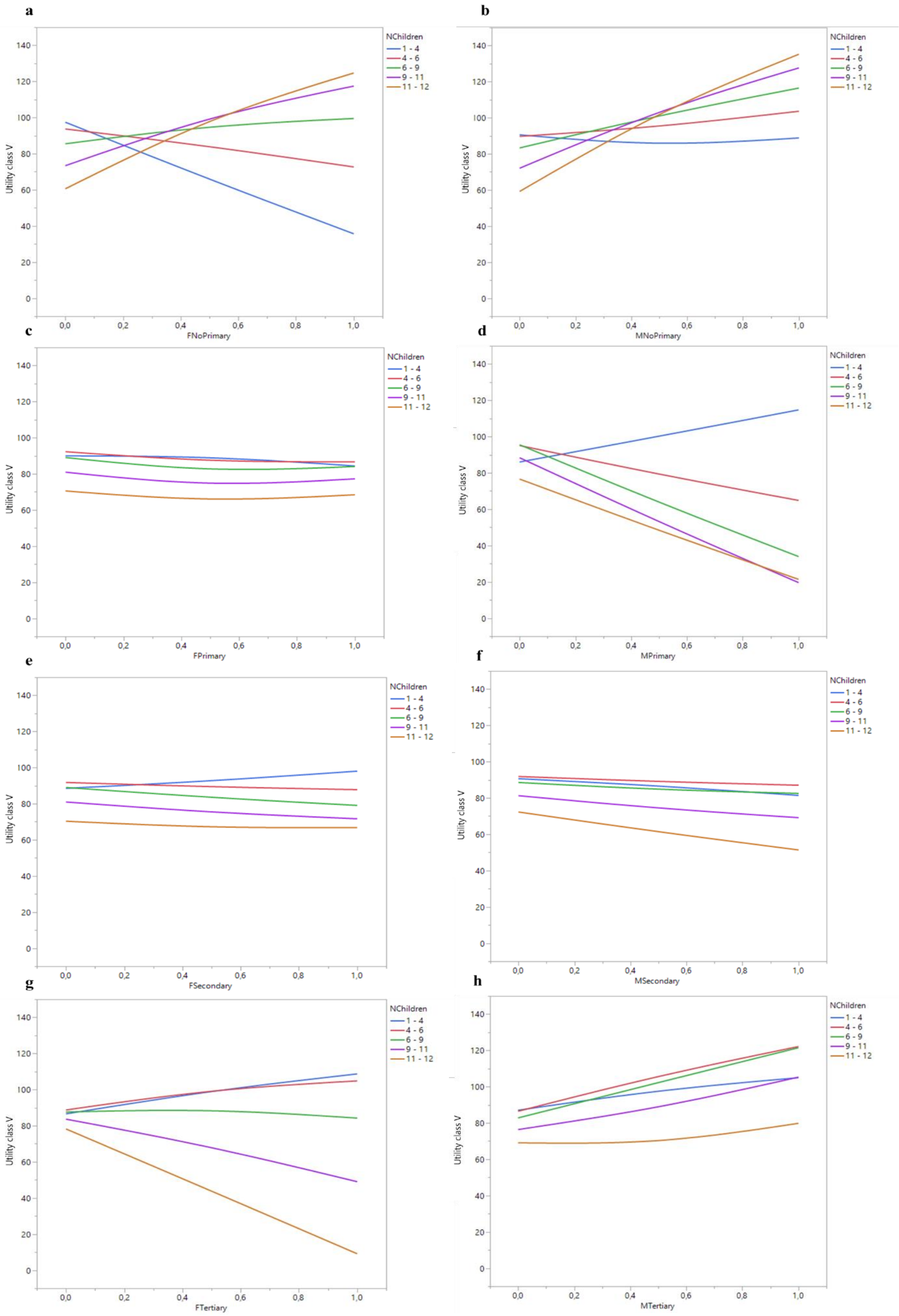


*Figure 8-10: Visualization of **LC V**'s utility in function of the different schooling components for different ranges of the number of children. Note: all results need to be interpreted with respect to FNoPrimary. The explanation of the variables can be found in Table 3-2.*

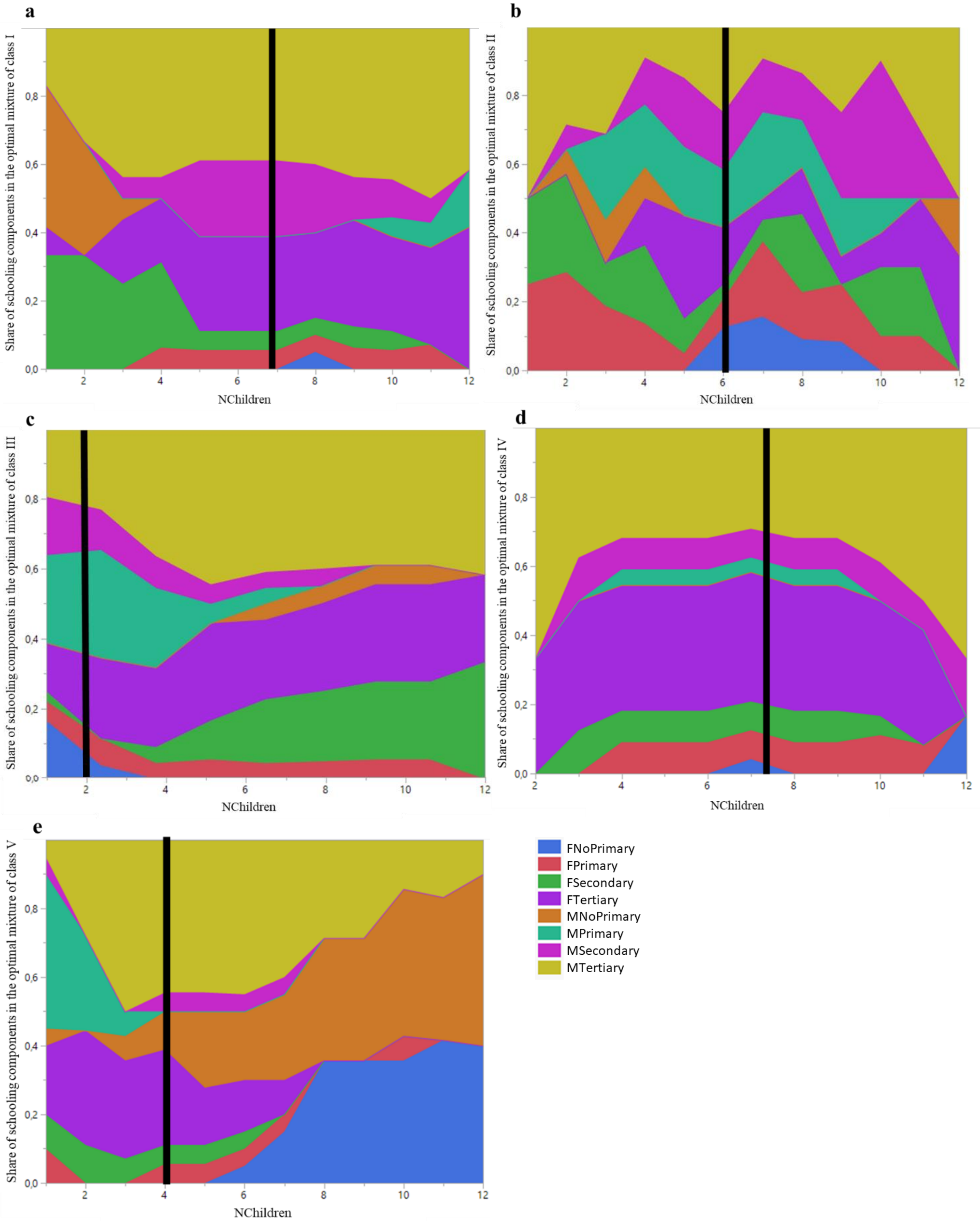


Figure 8-11: Optimal mixtures for the five LCs (a-e corresponding with LC I to V) for the schooling components as a function of the highest desirable number of children (indicated by the black line). Note: all results need to be interpreted with respect to FNoPrimary. The explanation of the variables can be found in Table 3-2.

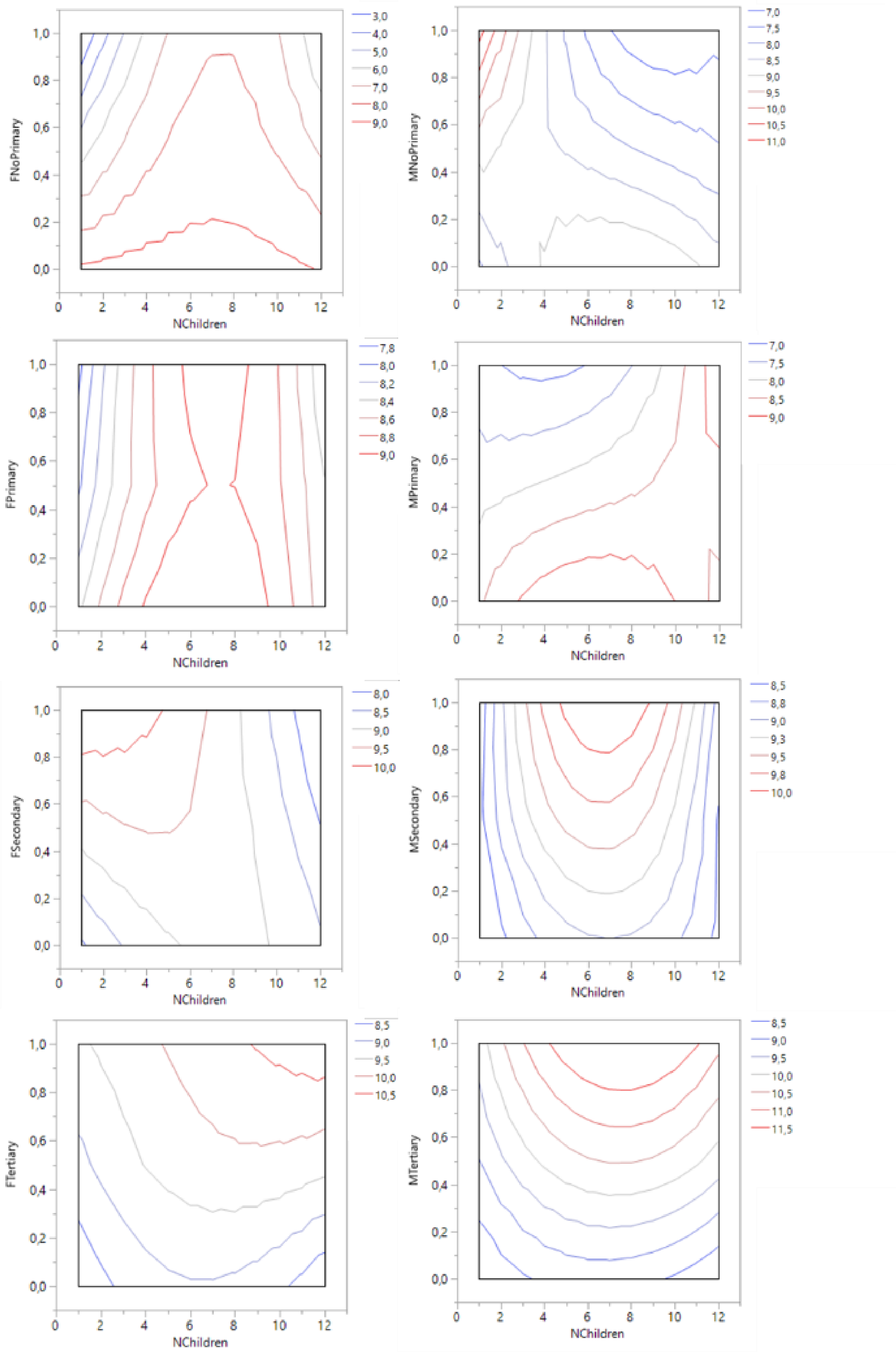


*Figure 8-12: Utility indifference curves for the different schooling components (y-axis) and the number of children (x-axis) to visualize the MRS for **LC I**. Note: all results need to be interpreted with respect to FNoPrimary. The explanation of the variables can be found in Table 3-2.*

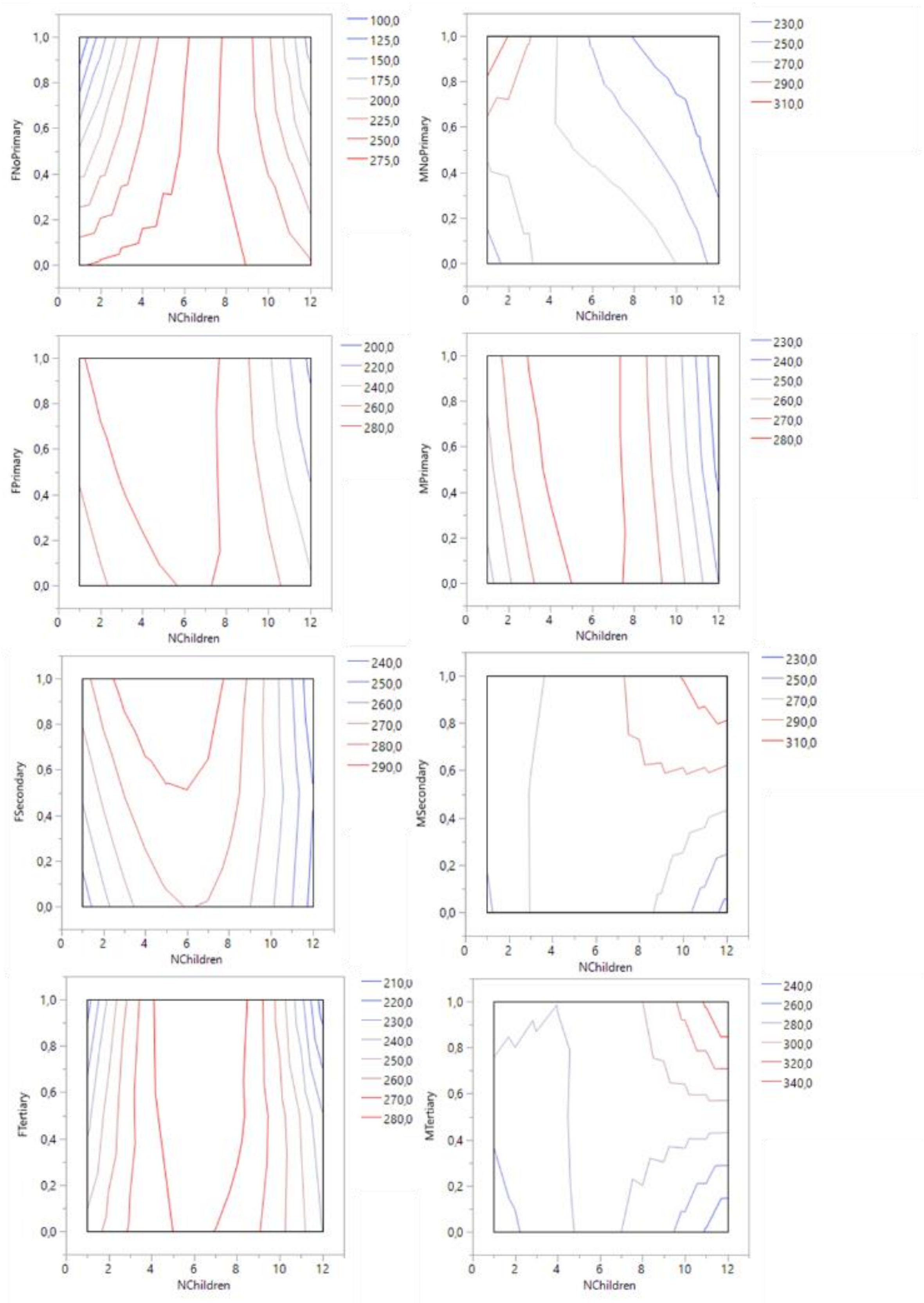


*Figure 8-13: Utility indifference curves for the different schooling components (y-axis) and the number of children (x-axis) to visualize the MRS for **LC II**. Note: all results need to be interpreted with respect to FNoPrimary. The explanation of the variables can be found in Table 3-2.*

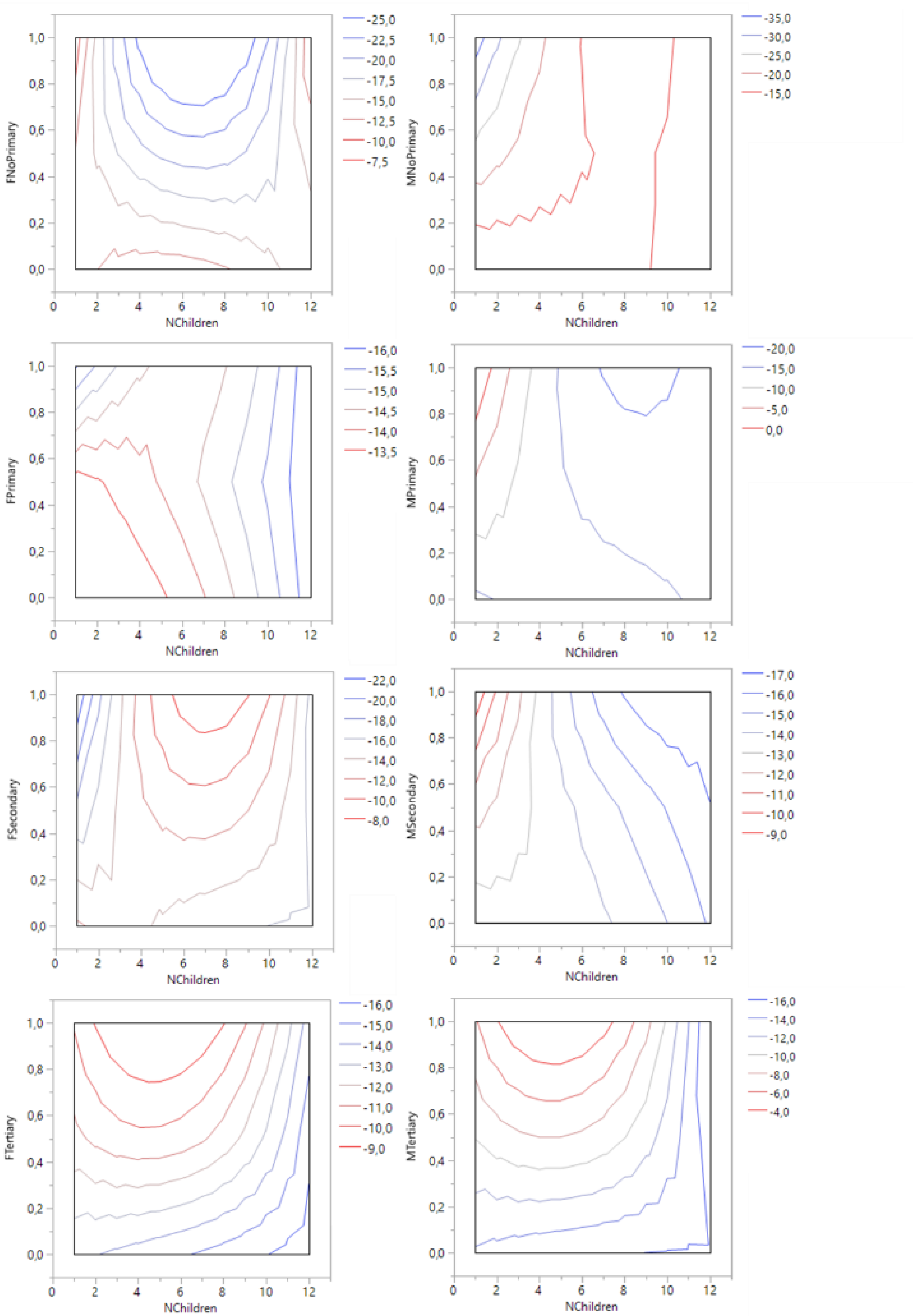


*Figure 8-14: Utility indifference curves for the different schooling components (y-axis) and the number of children (x-axis) to visualize the MRS for **LC III**. Note: all results need to be interpreted with respect to FNoPrimary. The explanation of the variables can be found in Table 3-2.*

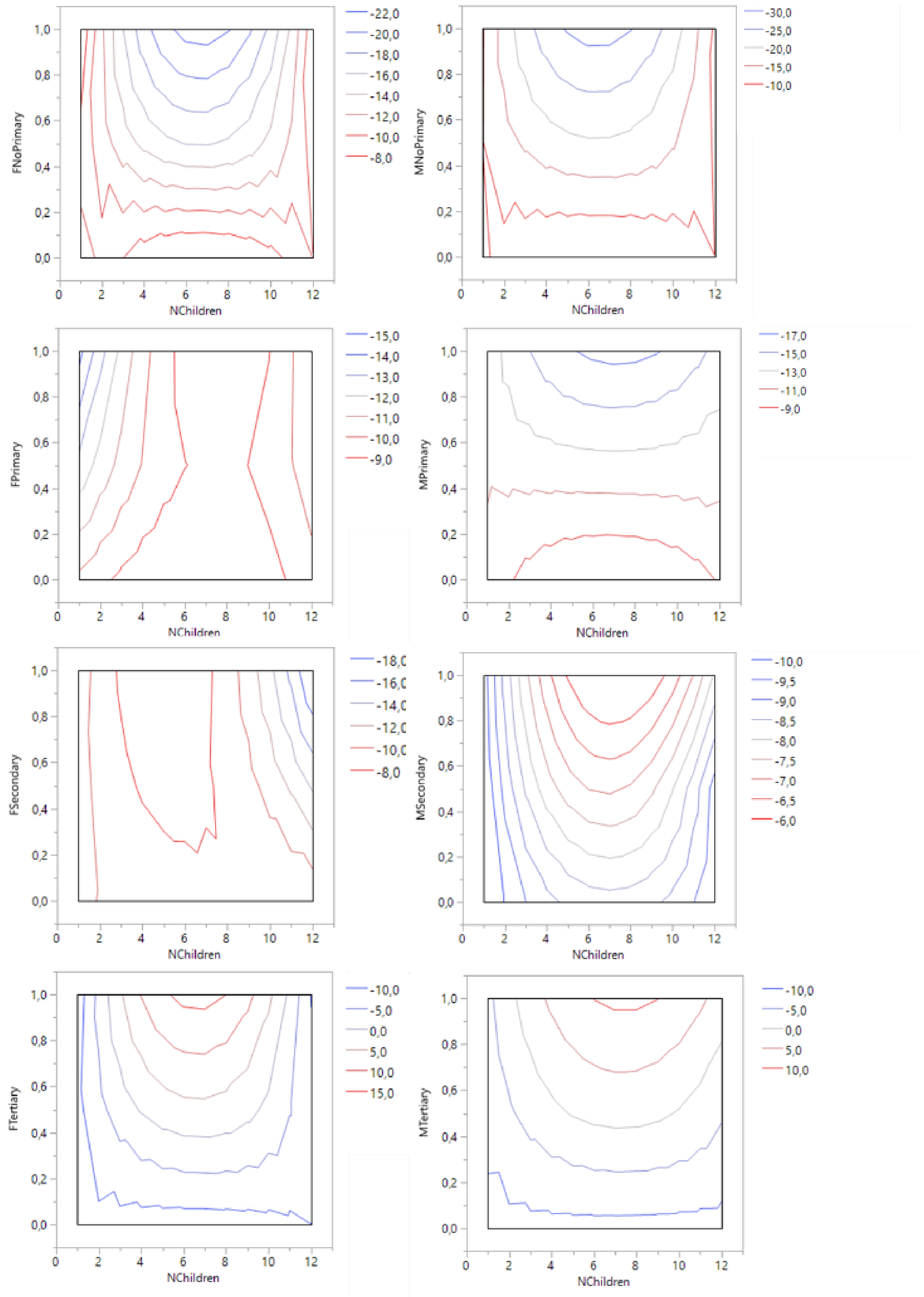


*Figure 8-15: Utility indifference curves for the different schooling components (y-axis) and the number of children (x-axis) to visualize the MRS for **LC IV**. Note: all results need to be interpreted with respect to FNoPrimary. The explanation of the variables can be found in Table 3-2.*

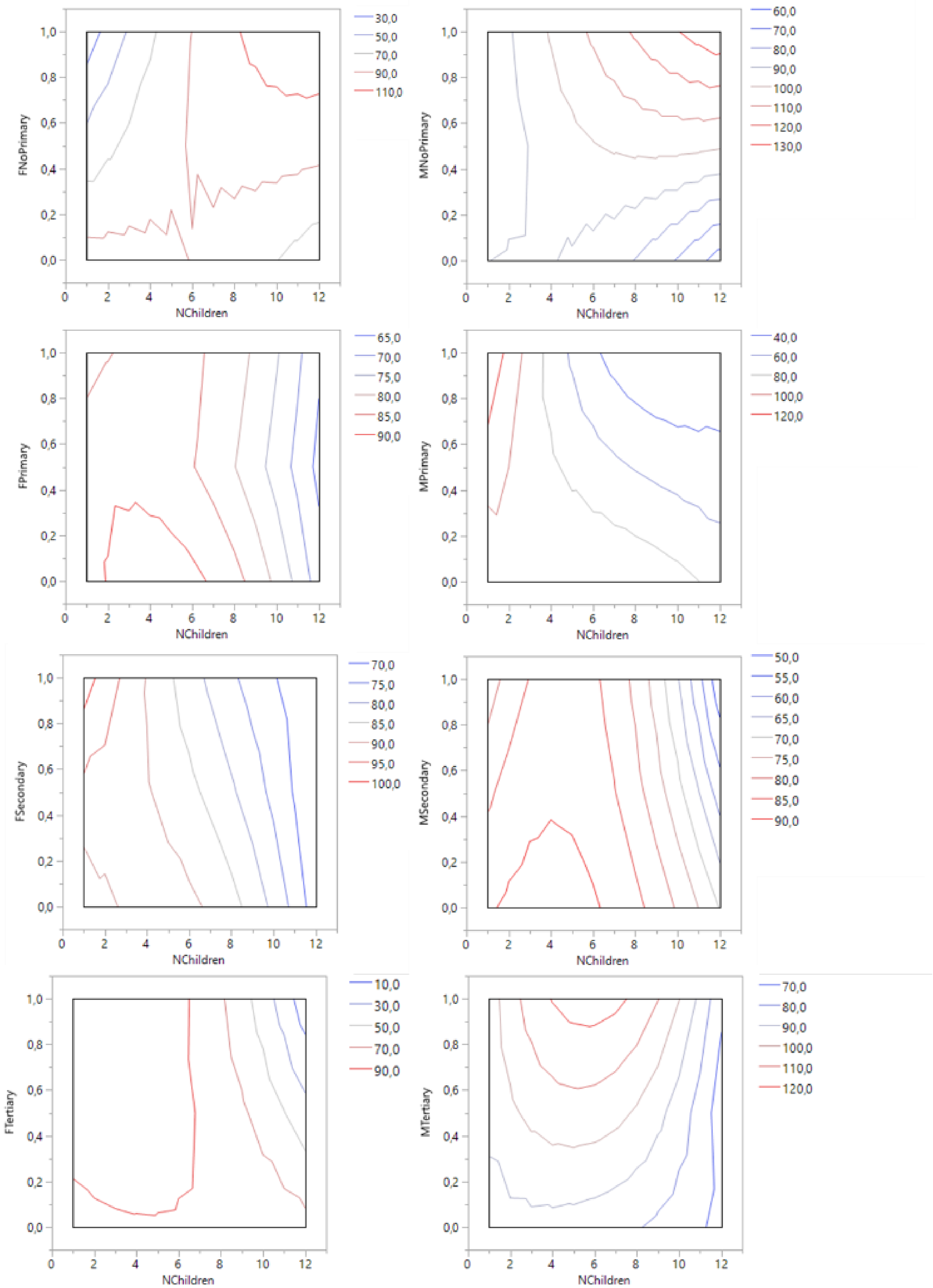


*Figure 8-16: Utility indifference curves for the different schooling components (y-axis) and the number of children (x-axis) to visualize the MRS for **LC V**. Note: all results need to be interpreted with respect to FNoPrimary. The explanation of the variables can be found in Table 3-2.*

9. Popularized summary

Since the start of the industrial revolution, the world population started to increase tremendously. Population growth is known to be influenced by fertility and mortality. Mortality rates drop due to scientific and technological progress, which is in general a sign of development that one does not want to slowdown for the sake of limiting population growth. Consequently, the major possible control on the current population growth is through fertility control. Future population predictions vary widely. Although 60% of the current world population lives in Asia, only the African continent is expected to increase in population the coming century. Therefore, it is necessary to study which factors are related with high fertility preferences in Africa. In addition, previous studies show a correlation between low parental education and high fertility. Consequently, it is important to investigate if high fertility rates and low schooling rates go hand in hand, since this can complicate the achievement of lower fertility rates. Further, girls are generally less educated than boys in Africa. Therefore, it is necessary to study whether the relation between high fertility and low schooling rates is different for boys and girls, since lower female education will lead to lower female empowerment in future generations which will again lead to more children.

In this thesis, an empirical case study was performed in Ethiopia. A choice experiment was conducted, followed by quantitative survey questions and some focus group discussions. Our study is innovative since a choice experiment with a non-linear mixture-amount model was applied. Our results have different policy implications. First, the general fertility preferences are on average 5.98 for men and 5.62 for women. It is recommended that family planning policies mainly focus on decreasing the prevailing family size norm next to the current focus on access to contraception. The following groups which have a higher prevailing ideal family size need to be targeted: men, people who are married, have children, are living in the Gamo Gofa Zone or Konso district and belong to the Konso ethnic group. This can be done for example through mass media, health centres, education, etc. Further, improvements in birth registration are needed to make national data more reliable. Second, it was shown that, since resources are constraint, people make a trade-off between the number of children they want and the investment in schooling. In general, this holds until the preferred number of children is reached. The existence of the trade-off implies that is possible to stimulate declines in quantity by increasing the availability, accessibility and affordability of education. Third, in general, men have a sex preference for sons concerning the preference for number of children and investment in education. To reduce the gender inequality in education, mainly men need to be targeted.