Connected Choices:

Marriage and Fertility in the Age of Social and Digital Networks in sub-Saharan Africa

Jet Wildeman

Institute for Management Research

RADBOUD UNIVERSITY PRESS

Radboud Dissertation Series

Connected Choices: Marriage and Fertility in the Age of Social and Digital Networks in sub-Saharan Africa

Jet Wildeman

Connected Choices: Marriage and Fertility in the Age of Social and Digital Networks in sub-Saharan Africa
Jet Wildeman

Radboud Dissertation Series
ISSN: 2950-2772 (Online); 2950-2780 (Print)

Published by RADBOUD UNIVERSITY PRESS
Postbus 9100, 6500 HA Nijmegen, The Netherlands
www.radbouduniversitypress.nl

Design: Proefschrift AIO | Guus Gijben
Cover: Proefschrift AIO | Guntra Laivacuma
Printing: DPN Rikken/Pumbo

ISBN: 9789465152493
DOI: 10.54195/9789465152493
Free download at: https://doi.org/10.54195/9789465152493

RADBOUD
UNIVERSITY
PRESS

Connected Choices: Marriage and Fertility in the Age of Social and Digital Networks in sub-Saharan Africa

Proefschrift ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.M. Sanders,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op

dinsdag 9 juni 2026
om 12.30 uur precies

door

Jet Wildeman
geboren op 7 december 1993
te Raalte

Promotoren:
Prof. dr. N. Wagner
Prof. dr. J.P.J.M. Smits

Manuscriptcommissie:
Prof. dr. C.H.B.M. Spierings
Prof. dr. R. Hartwig (Ruhr-Universität Bochum, Duitsland)
Dr. I. Bonfrer (Erasmus Universiteit Rotterdam)

Table of Contents

Chapter 1

1.1 Introduction

Sub-Saharan Africa (SSA) stands out in global demographic trends. While much of the global debate now centers on declining fertility in high-income countries (Goldin, 2025), SSA displays the opposite pattern, with persistently high fertility and early family formation. The average total fertility rate (TFR) in the region exceeds 4.3 children per woman, with some countries still reporting rates above 6, while the global average is 2.2 (United Nations, 2023). Moreover, nine of the ten countries with the highest prevalence of child marriage are located in SSA, and in 2023 the early adolescent birth rate (ages 10–14) was nearly three times the global average (Child Marriage Data Portal, 2024; UNICEF, 2015; WHO, 2024). These patterns carry far-reaching implications: they are strongly associated with lower educational attainment, poor maternal and child health outcomes, limited labor market participation, and intergenerational cycles of poverty (e.g. Chari et al., 2017; Field & Ambrus, 2008; Ganchimeg et al., 2014; UNFPA, 2012). Understanding what factors contribute to the persistence of early marriage, childbearing, and large family sizes is therefore essential, not only for improving individual well-being, but also for advancing broader development goals, such as ensuring universal access to reproductive health services (SDG 3, Target 3.7) and eliminating harmful practices like child marriage (SDG 5, Target 5.3).

High fertility, child marriage, and teen pregnancy rates in SSA are not only the outcomes of poverty or limited access to education and contraception, but are also deeply shaped by local social norms and cultural traditions (Bongaarts & Casterline, 2013; Caldwell & Caldwell, 1987; Casterline, 2017). Expectations around family size, the timing of marriage, and early childbearing are reinforced through ethnic identity, religious affiliation, and customary practices (Ahonsi et al., 2019; Kane et al., 2016; Schaffnit et al., 2019). These norms often define what is considered acceptable or desirable behavior for girls and young women, linking early marriage and motherhood to social recognition, respectability, and future security (Field & Ambrus, 2008; Ganchimeg et al., 2014; Jensen & Thornton, 2003). Reproductive preferences and behaviors are socially embedded and maintained through kinship ties and the intergenerational transmission of expectations (Johnson-Hanks, 2005; Trinitapoli & Yeatman, 2018). This helps explain why early family formation persists, even in the face of rising educational attainment and targeted policy interventions to counteract it (Batyra et al., 2021; Feyissa et al., 2023).

In this dissertation, I examine how two distinct forces - social norms embedded in group identities and access to digital information - shape women's family formation

in SSA. I focus on how expectations within ethnic, religious, and local communities influence women's decisions, and how digital technologies, such as social media and mobile internet, provide channels through which external ideas and role models reach them. Social groups shape behavior through shared norms, rewards, and sanctions, while digital technologies enable exposure to alternative norms and preferences beyond the immediate community. In this way, they can reinforce local expectations or open space for change by broadening the set of reference groups. By situating women within both their normative and informational environments, this dissertation offers a more contextualized understanding of early marriage, childbearing, and fertility preferences in SSA.

This dissertation is organized around two thematic blocks. In the first block, I focus on how the social context shapes women's family formation, with a focus on fertility preferences, the timing of marriage, and the onset of childbearing. I examine the ways in which individuals are embedded within normative environments, defined by ethnic, religious, or local group identities, and how these environments influence family formation beyond individual socioeconomic characteristics. A key contribution of this block is the development of a novel proxy for social pressure, based on within-group variation in family formation patterns. This measure provides an empirical strategy for capturing the strength of social pressure across different contexts in SSA.

The second block explores the role of digital connectivity in changing women's family formation trajectories. As social media and mobile internet use expand rapidly across SSA, they provide channels through which women gain access to new sources of information, alternative role models, and diverse social norms. It is not the technologies themselves that drive change, but the ideas and influences they transmit from beyond women's immediate communities. By using digital trace data from a major social media platform and geo-referenced 3G (mobile internet) coverage, combined with nationally representative demographic survey data, I assess how increased digital connectivity relates to changes in fertility behavior, child marriage, and teen pregnancy.

In the remainer of this chapter, I will discuss the theoretical framework (1.2). Next, I will discuss the four contribution I make in this dissertation (1.3), and I will provide an overview of the empirical chapters (1.4).

1.2 Theoretical background

1.2.1 The importance of social norms in marriage and fertility behavior

A growing body of literature emphasizes the importance of social embeddedness in shaping reproductive behavior, particularly in low- and middle-income settings (Batyra et al., 2021; Bursztyn & Jensen, 2017; Feyissa et al., 2023; Granovetter, 1985; Trinitapoli & Yeatman, 2018). Social embeddedness refers to the idea that individuals do not make decisions in isolation, but rather within a network of social relations, shared norms, and group expectations (Granovetter, 1985). Central to this embeddedness are social norms - informal rules that govern acceptable behavior within a group or society. These norms are learned through interaction, enforced through sanctions or rewards, and sustained through repetition and group approval (Bicchieri, 2006; Elster, 1989). In the context of family formation, women's choices around marriage and childbearing are often shaped by the expectations of kinship networks, ethnic and religious communities, and neighborhoods that define socially acceptable behaviors and impose sanctions for deviance (Ahonsi et al., 2019; Mkhwanazi, 2010). These group-level influences operate through both social pressure - such as reputational concerns or fear of exclusion - and social support, such as access to resources or protection, which are often conditional on conforming to group norms (Bernardi, 2003; Bourdieu, 1986).

This embeddedness is especially powerful in decisions that are highly visible and socially symbolic, such as marriage and childbearing (Akerlof & Kranton, 2000; Bernardi, 2003). In many parts of SSA, these transitions mark critical social milestones and serve as markers of adulthood, moral propriety, and group belonging (Ahonsi et al., 2019; Mkhwanazi, 2010; Schaffnit et al., 2019). As a result, women's decisions around family formation are often evaluated not only by their personal desires or circumstances but by how well they align with community expectations. This resonates with theories of social image utility, where individuals derive utility from being perceived positively by members of their reference groups (Bursztyn & Jensen, 2017). Social norms around the ideal timing and sequencing of marriage and fertility are often reinforced by community-level practices and shared histories, reducing the scope for autonomous decision-making.

Social embeddedness is particularly relevant in SSA due to the strength of informal institutions and community-based norm systems. Across the region, formal laws and state-based interventions- such as minimum age-of-marriage legislation - often coexist with, or are overridden by, traditional or religious authorities that

govern family life (Casterline, 2017; A. L. Robinson, 2014). These institutions operate at the local level, producing significant variation in behavioral norms across ethnic, religious, and regional lines. For example, studies have shown that marriage norms and fertility expectations differ widely between ethnic groups even within the same country, and these differences persist over time and across generations (Ahonsi et al., 2019; Schaffnit et al., 2019).

Moreover, classic individual-level predictors such as education or household wealth do not fully explain variation in family formation behaviors in SSA (Bongaarts & Casterline, 2013; Caldwell & Caldwell, 1987; Casterline, 2017). While these factors matter, their effects are often mediated or conditioned by the surrounding normative environment (Batyra et al., 2021; Trinitapoli & Yeatman, 2018). For instance, the impact of education on delaying marriage or reducing fertility may be limited if prevailing community norms prioritize early marriage or high fertility as symbols of female virtue or status. In such cases, women may experience competing incentives, where aligning with group norms ensures social approval and access to support, while deviation, despite potential long-term gains, may entail substantial personal or social cost (Field & Ambrus, 2008; Schaffnit et al., 2019).

1.2.2 New opportunities by digital technologies

In recent years, the diffusion of digital technology, particularly mobile internet and social media, offers the opportunity to shift the normative landscape in SSA (Fatehkia et al., 2018; Manacorda & Tesei, 2020; Rotondi et al., 2020). While traditional social norms are often maintained through face-to-face interaction and dense local networks, access to the internet enables exposure to alternative values, behaviors, and role models, many of which may stand in contrast to prevailing community expectations around marriage and fertility (Arias, 2019; Ferrara et al., 2012; Kearney & Levine, 2015). This creates a tension between local embeddedness and digital connectivity, whereby women and girls may come to question or renegotiate social norms that previously went unchallenged (Abubakar & Dasuki, 2018; Rotondi et al., 2020).

Theoretically, this process reflects a form of norm disruption through expanded reference groups and new channels of social influence(Bicchieri, 2006; Kietzmann et al., 2011; Rainie & Wellman, 2012). Whereas traditional norms are reinforced by immediate family, elders, and religious leaders, digital platforms such as Facebook, WhatsApp, and mobile search engines allow adolescents and young women to interact with peers across social and geographic boundaries. These platforms facilitate social learning, the spread of information about contraception

and education, and the emergence of new aspirations regarding life trajectories (Kearney & Levine, 2015; Rotondi et al., 2020).

Importantly, access to new information through digital technology alters the decision-making environment by reducing the information monopoly of the local community. In settings where social norms have long been transmitted and enforced through tight-knit groups, the ability to access diverse viewpoints, whether about reproductive health, marriage alternatives, or educational opportunities, can challenge the existing norms (Bicchieri, 2006). Even passive exposure to different life scripts, for example via online media or peer-generated content, may shift perceptions of what is desirable or feasible, especially for adolescents and young adults navigating critical transitions (Abubakar & Dasuki, 2018; Ferrara et al., 2012).

This informational shift is especially significant in contexts where formal sexuality education is limited or fragmented, and where intergenerational dialogue about reproductive autonomy is constrained by taboos or patriarchal norms (Guldi & Herbst, 2017; UNESCO, 2021). Digital connectivity offers not only the possibility to increase factual knowledge (e.g., about contraception or school continuation), but also introduces alternative framings of womanhood, adult roles, and social status. These framings can subtly or directly undermine traditional expectations that associate femininity with early marriage and childbearing (Rotondi et al., 2020; Varriale et al., 2022).

As a result, mobile internet and social media may contribute to normative pluralism - where multiple, sometimes conflicting, social scripts are simultaneously available to individuals. In this environment, women become embedded not only in immediate social networks that transmit local norms but also in virtual spaces where norms may be more fluid, aspirational, or individually tailored (Kietzmann et al., 2011; Rainie & Wellman, 2012). This layered embeddedness creates room for both reinforcement and contestation of existing expectations, depending on factors such as digital access, literacy, social position, and the salience of competing reference groups (Bongaarts, 2017; DiMaggio, Hargittai, Russell Neuman, et al., 2001).

1.3 Research questions

Building on the theoretical framework, this dissertation addresses the following research questions:

1. To what extent and in what ways do fertility preferences in SSA vary between ethnic groups? And how can this variation be explained by socioeconomic, demographic, and cultural characteristics of ethnic groups?
2. How does the strength of social norms, as captured by variation in the timing of marriage and first birth within social groups, affect young women's incidence of early marriage and adolescent childbearing in SSA?
3. To what extent is regional variation in crude birth rates associated with levels of social media usage and gender gaps in digital access across SSA?
4. How does the expansion of mobile broadband (3G) coverage influence the incidence of child marriage and teen pregnancy in SSA?

1.4 Data and methods

1.4.1 Data

The analyses in this dissertation are based on a combination of household survey data, digital trace data, and geospatial information. The main source is the Demographic and Health Surveys (DHS; https://dhsprogram.com/), complemented where relevant by the Multiple Indicator Cluster Surveys (MICS; https://mics.unicef.org/). These nationally representative surveys, conducted in dozens of sub-Saharan African countries since the 1980s, provide detailed information on fertility, marriage, education, and socioeconomic background. Their standardized design and wide coverage across time and space allow for broad comparative analyses across countries, ethnic groups, and cohorts.

To examine the role of digital connectivity, these surveys are complemented with two additional data sources. Subnational digital trace data from Facebook's Marketing API (https://developers.facebook.com/docs/marketing-api) provide indicators of social media usage and gender gaps in online access. In addition, georeferenced maps of mobile broadband coverage from the GSM Association (GSMA) and Collins Bartholomew Mobile Coverage Maps (https://www.collinsbartholomew.com/coverage-mapping/) supply annual information on the expansion of 3G networks across Africa. Combining these datasets makes it possible

to link women's family formation trajectories to both their immediate normative environment and their exposure to new digital infrastructures.

Where relevant, these data are further enriched with supporting sources such as World Bank development indicators (https://data.worldbank.org/), Afrobarometer surveys (https://www.afrobarometer.org/), and geospatial measures of for example rainfall and lightning. These complementary datasets are discussed in detail in the empirical chapters where they are applied.

1.4.2 Methods

The empirical chapters apply a range of quantitative methods suited to the different research questions. Variation in fertility preferences across ethnic groups is analyzed with multilevel Poisson regression models, which account for both individual and group-level influences. To capture the strength of social pressure on marriage and childbirth timing, I use fixed effects regressions that incorporate a novel proxy based on within-group variation in family formation behavior. The relationship between social media usage and fertility outcomes is examined through spatial analysis and regression models linking Facebook activity to subnational birth rates while controlling for development indicators. Finally, the impact of mobile internet expansion on child marriage and teen pregnancy is estimated with quasi-experimental strategies, including staggered difference-in-differences, event studies, and an instrumental variables approach.

Together, these methods combine descriptive and causal inference strategies. They make it possible to study how deeply embedded social norms and rapidly expanding digital connectivity interact in shaping women's family formation trajectories in sub-Saharan Africa.

1.5 Key contributions

In this dissertation, I make four key contributions to the literature on fertility, family formation, and social change in SSA, particularly within the framework of development economics, demography and sociology. I introduce empirical insights into how the social context, cultural norms, and digital technologies enable channels to shape marriage and fertility behavior.

A first contribution lies in foregrounding ethnicity as a central axis of variation in fertility preferences across African contexts. While previous research has often

treated ethnicity as a control variable or a proxy for culture, this dissertation places it at the analytical core. By emphasizing how ethnic group norms structure family size preferences, it responds to longstanding calls in demography to better integrate culture into models of fertility (Bongaarts & Watkins, 1996; Johnson-Hanks, 2007). It advances the view that cultural identity, expressed among others through group-level gender norms and educational patterns, mediates the relationship between individual characteristics and fertility preferences, thereby challenging purely individualistic or economic explanations.

Second, this dissertation develops a novel, behavior-based proxy for social pressure by measuring within-group variation in the timing of marriage and childbirth. This builds on social utility models that links norm strength to behavioral conformity (Bursztyn & Jensen, 2017), while addressing a persistent empirical challenge: the lack of scalable, comparable indicators of normative influence in large-scale datasets. By linking uniform behavior to strong norms, this method offers a clear and workable way to capture social pressure. This approach advances efforts to systematically capture how social environments and normative pressures shape family formation behavior, providing a scalable way to study their influence on family formation.

A third contribution is the integration of digital connectivity, particularly social media, as a channel of change for fertility rates. While prior research has documented the influence of traditional media on fertility preferences (e.g. Ferrara et al., 2012; Jensen & Oster, 2009), this dissertation brings new attention to the role of interactive, user-driven platforms in shaping family formation behavior. It demonstrates that digital exposure is related to lower fertility, possibly because digital exposure serves as a conduit for alternative scripts around gender, reproduction and life aspirations (Fatehkia et al., 2018; Rotondi et al., 2020). This work also contributes methodologically by incorporating digital trace data into fertility research, offering a novel lens through which to examine subnational variation in demographic outcomes.

Finally, the dissertation adds to the causal literature on fertility and marriage by providing evidence on the impact of mobile broadband expansion on early family formation outcomes. By leveraging exogenous variation in 3G rollout and applying quasi-experimental designs (e.g., difference-in-differences and instrumental variables), it contributes to the literature on ICT and social development (Manacorda & Tesei, 2020). This work illustrates that mobile internet can change social behaviors. Importantly, it shows that these impacts are shaped by existing inequalities, pointing to the importance of assessing who gains and who is left behind.

Together, these contributions offer a multidimensional perspective on marriage and fertility behavior that links enduring social structures with rapidly evolving digital environments. Theoretically, the dissertation synthesizes insights from cultural demography, social norm theory, and models of informational influence. Methodologically, it develops scalable tools for capturing social pressure and leverages geospatial and digital data to generate new empirical leverage. From a policy standpoint, it provides a foundation for more context-sensitive, norm-aware, and digitally inclusive interventions to support adolescent reproductive health and gender equity. In doing so, it responds to recent calls in development economics and public health to move beyond individual-level approaches and engage more fully with the social contexts in which decisions about marriage and fertility are made (Bongaarts & Watkins, 1996; Bursztyn & Jensen, 2017; Johnson-Hanks, 2005; Trinitapoli & Yeatman, 2018).

1.6 Chapter overview

1.6.1 Chapter 2 – Ethnicity and family size preferences

This chapter examines how fertility preferences vary across ethnic groups in SSA, using data from over 500,000 women across 181 ethnic groups in 24 countries. It explores the role of ethnicity as a cultural framework influencing preferred family size, going beyond socioeconomic indicators like education and wealth. Using Multilevel Poisson models shows that substantial variation in preferred family size is coming from a woman's ethnic background.

While part of this variation is explained by compositional factors (e.g., place of living, education), ethnicity remains a strong predictor after controlling for individual and household characteristics. Ethnic group-level characteristics such as educational attainment and gender inequality also influence fertility preferences, with women in more gender-equal or better-educated groups preferring fewer children. The interaction analysis suggests that women's individual characteristics, like their level of education, are more impactful in cultural contexts characterized by higher female empowerment and weaker enforcement of traditional gender norms.

These findings underscore that ethnic identity shapes expectations around fertility, and that this influence persists even when education and living standards improve. This chapter lays the foundation for understanding how localized cultural norms affect fertility behavior and introduces the idea that group-level social environments shape the conditions under which women can exercise reproductive agency.

1.6.2 Chapter 3 – Social pressure and family formation

This chapter investigates the role of social pressure on women's timing of marriage and childbirth. Using 104 DHS surveys across 33 countries, it introduces a novel proxy for measuring normative pressure: the coefficient of variation in age at first marriage and first birth within women's local, ethnic, and religious groups.

Using fixed effect regression analysis, I find that greater within-group variation - interpreted as weaker norms - is linked to later marriage and childbirth and a lower likelihood of child marriage and teen pregnancy. Specifically, a one standard deviation increase in variation delays family formation by 4–5 months and reduces early transitions by nearly 2 percentage points. These effects are strongest when the reference group is similar in age, suggesting that peer norms are especially influential for fertility, while the timing of marriage is also shaped by older generations.

This approach contributes a scalable method to quantify social pressure and highlights the persistence of local norms despite broader development trends. The findings suggest that delayed family formation is not only about access to education or resources but also about the degree of behavioral flexibility allowed within social groups.

The chapter advances the argument that understanding variation in family formation is critical to addressing early transitions. It also provides a conceptual bridge to the next chapters, which explore how digital information environments - such as social media and mobile internet - might reshape these norms.

1.6.3 Chapter 4 – Social media and birth rates

This chapter examines how social media usage relates to fertility outcomes across 311 subnational regions in 29 SSA countries. Using digital trace data from Facebook combined with household survey data from the DHS, it investigates whether higher levels of social media engagement - and greater gender equality in usage - correlate with lower crude birth rates.

Findings show a robust negative association between social media usage and fertility rates, even after accounting for education, wealth, and urbanization. Regions with high Facebook activity and narrower gender gaps in use exhibit significantly lower birth rates. These patterns are in line with the idea that social media facilitates the diffusion of modern reproductive norms.

The chapter builds on diffusion and media influence theories, arguing that social media provides not only access to information but also opportunities for identity construction and norm negotiation beyond the immediate community. It also highlights how gender inequality in digital access sustains higher fertility, while inclusive digital environments are linked to lower fertility rates.

By showing how virtual networks may influence real-life demographic outcomes, the chapter contributes new insights into how fertility behavior is formed and challenged. It complements the previous chapter on social pressure by demonstrating that the online environment can be a space where alternative norms take root - particularly when women are active participants in digital life.

1.6.4 Chapter 5 – Mobile internet and child marriage and teen pregnancy

This chapter evaluates how mobile broadband (3G) expansion affects child marriage and teen pregnancy in SSA. Using geo-referenced 3G coverage data linked to over 500,000 women-year observations from 20 countries, it applies causal models - including staggered difference-in-difference studies and instrumental variables - to estimate the effects of internet access.

Results show that adolescent girls exposed to 3G are significantly less likely to marry early or have a teen birth. Depending on the model, exposure to 3G reduces child marriage by 1–6 percentage points and teen pregnancy by about 1 point. These effects are especially pronounced among more educated and wealthier girls, suggesting differential adoption and benefit. Mechanism analysis supports the idea that access to information, contraception, and changed gender attitudes are key pathways.

This chapter highlights mobile internet as a tool that can challenge traditional social norms, particularly in settings where formal policies have had limited success. Unlike media exposure alone, mobile internet allows for interactive engagement and self-directed learning, giving young women greater autonomy over reproductive decisions.

By linking digital connectivity to measurable changes in early transitions, the chapter shows how new infrastructures can shift long-standing behaviors. It reinforces themes from earlier chapters - norm strength, peer influence, and information flows - by demonstrating that access to digital tools can reshape life trajectories, especially for young women navigating restrictive environments.

1

Chapter 2: Ethnic variation in fertility preferences in sub-Saharan Africa[1]

1. This chapter is based on joint work with Jeroen Smits and Sandor Schrijner and has been published as:

   Wildeman, J., Smits, J., and Schrijner, S. (2023). Ethnic variation in fertility preferences in Sub-Saharan Africa. *Population Research and Policy Review, 42-58 (2023).* https://doi.org/10.1007/s11113-023-09804-z.

   The study on which this chapter is based was presented at the Population Association of America Conference 2021 and won the best poster award.

Research focusing on the role of ethnicity in shaping fertility preferences in sub-Saharan Africa (SSA) has been largely restricted to small-scale studies within specific groups, regions or countries, which makes it difficult to gain insight into the overall effects of ethnicity on fertility in the region. This study provides a broad comparative analysis of the variation in preferred family size among ethnic groups covering the whole SSA subcontinent. The relationship between ethnicity and family size preferences is analyzed for 500,000 women from 181 ethnic groups in 24 sub-Saharan countries using data from the Demographic and Health Surveys and via multilevel Poisson regression analyses. Fertility preferences vary considerably across ethnic groups. Major factors explaining this are gender-related cultural characteristics and educational level at the group level. Interaction analyses showed that the size and direction of these effects are moderated by the women's individual characteristics. The findings make clear that someone's ethnic background should be taken into account when looking at fertility patterns and designing policies aimed at changing fertility patterns in the SSA context.

2.1 Introduction

Within Sub-Saharan Africa (SSA), fertility levels and a preference for large families are much higher than in most other regions of the developing world, even if we take socioeconomic position and the relatively low level of access to contraceptives into account (Bongaarts, 2011; Casterline, 2017; Dibaba & Mitike, 2016; United Nations, 2020). One possible explanation for this is that in SSA, more than elsewhere, fertility behavior depends on the socio-cultural context in which people live (Cole & Geist, 2021; Johnson-Hanks, 2005; Longwe & Smits, 2013; Trinitapoli & Yeatman, 2018). This context determines to a large extent what information and resources individuals have access to. It constructs leading norms and values, which may include powerful ideals about what constitutes appropriate reproductive behavior for women (Kane et al., 2016; Munshi & Myaux, 2006).

An important aspect of the social context in SSA is ethnicity. Ethnicity is a complex social construct encompassing the cultural aspects of life that influences individuals' identity and group social relations (Airhihenbuwa, 2007; Ford & Kelly, 2005). Members of ethnic groups can possess shared attributes, including beliefs, customs and/or shared memories and experiences. The level of ethnic diversity in SSA is higher than in most other global regions, and individuals' loyalty towards their ethnic group is often stronger than their loyalty towards their country (Bates, 2000; Casterline, 2017; Easterly & Levine, 1997; A. L. Robinson, 2014).

A number of empirical studies have shown that ethnicity plays a major role in women's fertility behavior. Addai (1999) found clear ethnic differentials in the use of contraceptives in Ghana. Kane et al. (2016) reported that women's fertility behavior within the Fertit community in South Sudan corresponded to the social norm of having as many children as possible. Palamuleni (2014) found differences in fertility rates between six ethnic groups in Malawi, and Munshi et al. (2006) concluded that changing social norms within ethnic groups were predominantly responsible for observed changes in reproductive behavior in rural Bangladesh. As such, available evidence suggests that ethnicity plays an important role in shaping women's fertility behaviors. However, not much is known about the strength of this influence, what characteristics of ethnic groups are important, and how these relate to fertility preferences.

The current preference for high fertility has important consequences for fertility levels in SSA, as the spread of a preference for smaller families has been one of the major factors responsible for fertility decline in other parts of the world

(Bongaarts, 2009; Coale & Watkins, 1986; Gerland et al., 2017). Given this preference for high fertility and its significance for any attempts to lower birth rates, gaining insight into the factors that influence it is of great importance. The current study focuses on the role of ethnicity in this respect, and aims to answer the following research questions:

1) *To what extent and in what ways do fertility preferences in SSA vary between ethnic groups?*

2) *How can this variation be explained by socioeconomic, demographic, and cultural characteristics of ethnic groups?*

Using data from the Demographic and Health Surveys (DHS; www.dhsprogram.com) with information on 500,000 women belonging to 181 different ethnic groups in 24 SSA countries, we study the relationship between ethnicity and family size preferences using multilevel Poisson regression models. In the next section, we provide an overview of the relevant literature regarding the relationships between ethnicity and fertility in the SSA.

2.2 Background

2.2.1 Theoretical background to fertility preferences in SSA

Preference for high fertility in SSA is pronounced compared to developing regions in other parts of the world, such as Latin America and Asia (Alaba et al., 2017; Bongaarts & Casterline, 2013; Channon & Harper, 2019). Although preferred family size has diminished somewhat over the years from above 6 in a majority of countries to 4.6 in 2010, it was still much larger then than in North Africa and the Middle East (3.2), Asia (2.7) and Latin America (2.7) (Bongaarts, 2011). An important reason for this preference lies in traditional pronatalist social, economic, and cultural practices (Bongaarts, 2011; Caldwell & Caldwell, 1987; Casterline, 2017; Hertrich, 2017). The pronatalist culture even implies a resistance to fertility decline that is absent in other developing regions (Bongaarts, 2017).

This preference for high fertility constitutes one factor, among others, that explains SSA's high fertility levels (Bongaarts, 2020; Bongaarts & Casterline, 2013; Ezeh et al., 2009; Westoff & Cross, 2006). Earlier studies have found that the spread of a preference for smaller families has been one of the major factors responsible for new fertility behavior and fertility decline in other parts of the world (Bongaarts, 2009;

Coale & Watkins, 1986; Easterlin & Crimmins, 1985; Gerland et al., 2017). Another important factor is the (lack of) access to family planning and related high unmet need for contraception (Cleland et al., 2006; Günther & Harttgen, 2016; Sedgh & Hussain, 2014). The lack of access to family planning can make it more difficult for Africans to translate smaller family size preference into practice (Günther & Harttgen, 2016). However, given that the preference for large families is still extraordinarily high, the relationship between this preference and the deeply embedded culture of loyalty towards ethnic groups seems to offer a highly interesting topic for empirical analysis (Bates, 2000; Casterline, 2017; Easterly & Levine, 1997; A. L. Robinson, 2014).

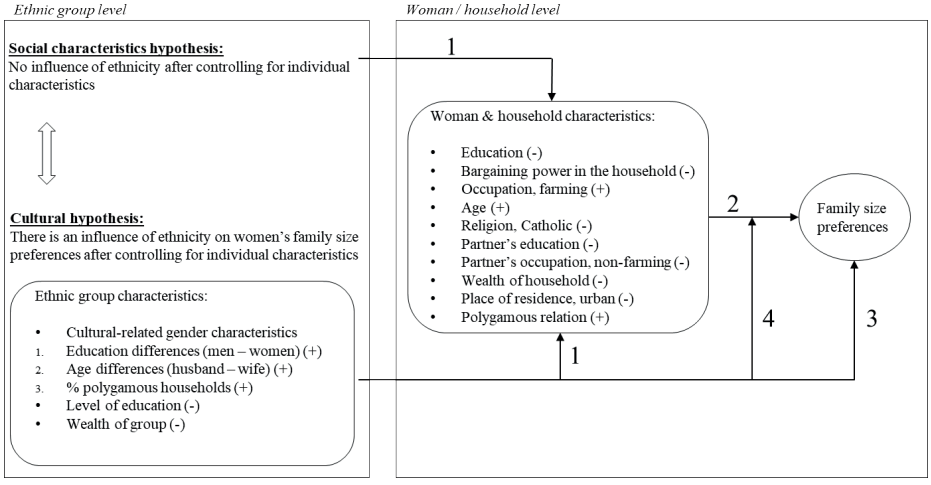
2.2.2 Ethnic fertility differentials

Attempts to understand ethnic fertility differentials have generally focused on two partly competing explanations: the social characteristics hypothesis and the cultural hypothesis. According to the s*ocial characteristics hypothesis* (Figure 2.1, arrow 1 and 2), ethnic fertility differentials are the result of compositional differences between groups, such as differences in socioeconomic, demographic, and cultural characteristics (Petersen, 1969). The basic assumption within this hypothesis is that individuals with similar characteristics may also show similar fertility preferences and behavior, irrespective of their ethnic background. Important characteristics in this respect are education (Shapiro & Tambashe, 1997), especially of women (Bakibinga et al., 2016; Kebede et al., 2019), traditionalism regarding gender roles (Hindin, 2000; Withers et al., 2015), and religion (Caldwell & Caldwell, 1987). The social characteristics hypothesis therefore predicts that ethnic fertility differentials will completely disappear when individual characteristics are taken into account. Empirical evidence in favor of this hypothesis is presented by Addai (1999) who found that ethnic differentials in contraceptive acceptance and use in Ghana disappeared when individual socioeconomic and demographic characteristics were controlled for, and by Bakibinga et al. (2016) who found that ethnicity was not associated with women's attitudes towards acceptance and use of family planning in Kenya.

The second hypothesis, the *cultural hypothesis*, assumes an additional influence of ethnicity on fertility behavior (Figure 2.1, arrow 1, 2 & 3) (Caldwell & Caldwell, 1987). This hypothesis acknowledges the importance of individual socioeconomic and demographic characteristics in explaining ethnic fertility differentials, but emphasizes that norms and values of (ethnic) groups are important additional characteristics that influence fertility behavior. The cultural hypothesis is supported by several studies showing that fertility differences between ethnic groups remain after controlling for the individuals' socioeconomic and demographic characteristics (Kollehlon, 2003; Palamuleni, 2014).

To be able to distinguish between the social characteristics hypothesis and the cultural hypothesis, individual characteristics that are known or expected to influence family size preferences should be taken into account to eliminate group compositional differences as far as possible. Important individual and household characteristics are women's education, their bargaining power within the household and their occupational resources (Atake & Gnakou Ali, 2019; Bakibinga et al., 2016; Kebede et al., 2019). Higher-educated women, women with more bargaining power and women with a non-farming occupation prefer fewer children. Other individual and household characteristics that are expected to be important are their partners' education and occupation, household wealth, place of residence, age, religion, and whether they practice polygamy (Atake & Gnakou Ali, 2019; Bongaarts, 2020; Kebede et al., 2019; Lerch, 2019; Matovu et al., 2017; Shapiro & Tenikue, 2017; Tomkinson, 2019). We expect that women prefer fewer children if they have partners with a higher level of education and a non-farming occupation, are wealthier, live in an urban setting, are younger, are Catholic and have monogamous relationships.

Figure 2.1 Conceptual framework of hypotheses tested in this study.



Notes: The different arrows denote the tested hypotheses. Note that arrow 1 shows the same effect: ethnicity is associated with woman and household characteristics. The social characteristics hypothesis predicts a pathway of ethnic influences through 1 & 2 but not 3. The cultural hypothesis predicts a pathway of ethnic influences through arrows 1, 2 & 3. We further test for moderations between ethnic group characteristics and individual characteristics with arrow 4. The signs after the included variables denote whether we expect a positive or negative influence.

2.2.3 Mechanisms

If norms and values at the ethnic group level influence women's fertility behavior, they may do so through socialization and/or social contagion. Socialization implies that individuals' values – including those related to fertility preferences – are formed at an early age and become deeply rooted in personality structure (Afulani & Asunka, 2015; Kolk, 2014). Social contagion implies that fertility patterns are copied from relevant others (Lois & Arránz Becker, 2014). Individuals living in the same social context, such as members of the same ethnic group, mutually influence each other's behavior (Bernardi & Klärner, 2014; Levy & Nail, 1993). As such, women belonging to the same ethnic group are likely to end up showing similar fertility behavior.

According to Berndt, Rodermund, and Timm (2018), two mechanisms are important for socialization and social contagion: social pressure and social support. Social pressure results from anticipated rewards and punishments for compliance or non-compliance with the expectations of the social context (Bernardi, 2003). These expectations do not need to be particularly overt. Simply observing the behavior and choices of others within the same group may lead to an urge to reduce cognitive dissonance between one's own and the group's behavior (Festinger, 1954). Individuals generally do not want to differ too much from the group they feel attached to. So, if women observe that other women within their ethnic community prefer large families, they are more likely to do so as well.

Social support refers to the advancement a woman may achieve from being a member of a social group. Social networks offer social capital, defined as resources embedded within social relationships (Bourdieu, 1986; Putnam, 1993). Social capital can take different forms, such as financial, practical, and emotional support from other members of the group, including the potential to meet the cost of having children. Social pressure and social support separately or together may stimulate women to exhibit (fertility) preferences that are in line with the group's expectations.

Women's fertility preferences may also be affected by gender-related cultural characteristics within groups. In groups where there are larger gender differences, women are often viewed as caregivers and as subservient to men, and their domain may be more strongly restricted to the private sphere (Gündüz-Hoşgör & Smits, 2008; Spierings et al., 2010). In such a context, having more children may be an important way of achieving social status (Casterline, 2017). As such, we predict that women prefer more children where there are larger gender differences (Figure 2.1, arrow 3).

Level of education of group members may also be highly relevant to fertility preferences (Kravdal, 2012). Families are generally smaller in groups where individuals are more highly educated (Bakibinga et al., 2016; Kebede et al., 2019). This could imply less social pressure to prefer large families, and more social pressure to go to school. Both forms of social pressure are likely to lead to women preferring smaller families.

2.2.4 Moderating effects

The effect of ethnic group characteristics cannot be expected to be the same for all women, as they are likely to be moderated by individual characteristics and the situation in which they are living (Longwe & Smits, 2012; Maralani, 2008). This raises the question as to whether women's individual socioeconomic resources can be strong enough to overcome group pressure (called by Spierings et al. (2010) "individual dominance"), or whether group pressure overrides individual characteristics. If the latter, women may feel they have to adjust their family size preferences to the group norm independent of their own personal resources ("situational dominance").

The situational dominance and individual dominance hypotheses will be tested in relation to both the level of education and gender-related cultural characteristics of ethnic groups, as both group features are expected to affect women's family size preferences. Both features will be interacted with three essential factors at the individual level – women's education, their bargaining power within the household, and their occupational resources – which are all known to affect family size preferences (Atake & Gnakou Ali, 2019; Bakibinga et al., 2016; Kebede et al., 2019).

2.3 Data and Methods

2.3.1 Data

For this study, data from the Demographic and Health Surveys (DHS; www.dhsprogram.com) were used. DHS are large, nationally representative household surveys. For each survey, non-overlapping areas (often enumeration areas) were randomly selected. These areas (called "clusters" henceforth) are usually communities, villages, or city quarters. In the selected clusters, all households were listed, and a random selection of 25–30 households were selected for interviews. First, household surveys were conducted in which basic information was collected about all household members, and subsequently all women aged 15 to 49 and usually resident were invited for an oral interview in which a large amount of information was obtained,

including in-depth data on socioeconomic, demographic, and (reproductive) health-related issues.

The inclusion criteria for surveys in this study were that 1) at least two DHS surveys were available for a specific country since 1998, and that 2) information on women's ethnicity was a feature of these surveys. The second criterium in particular meant that we were not able to include all countries or some of the most recent DHS surveys, as ethnicity-related questions were not always included.

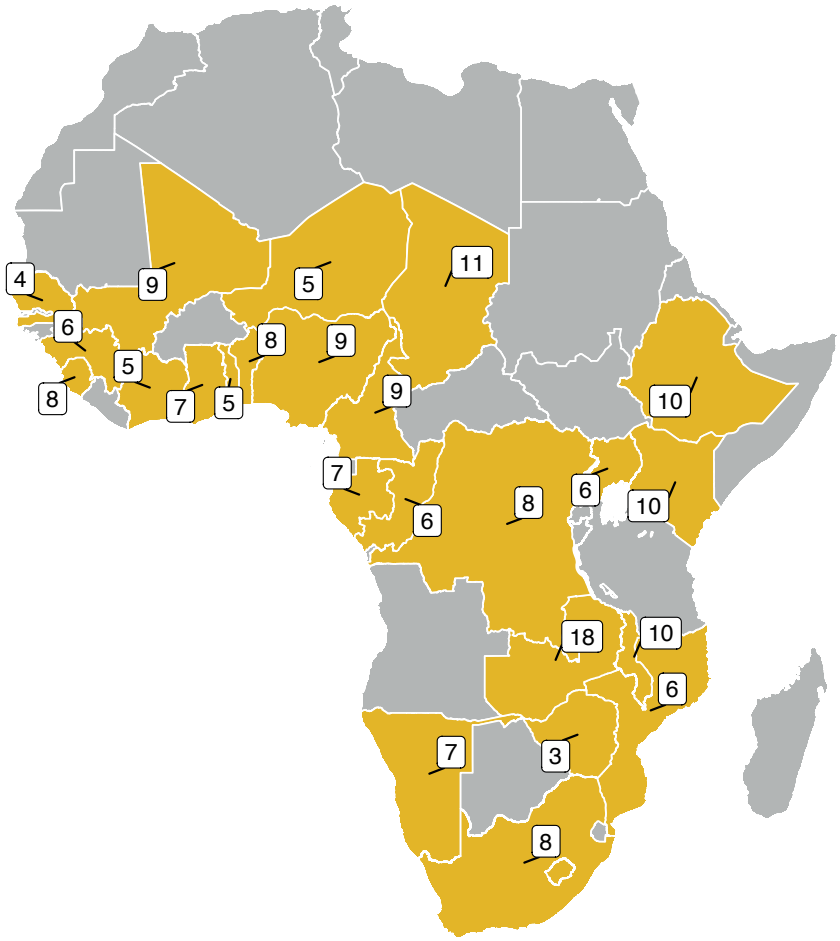
Information about women's ethnicity was asked in the household surveys. The question mostly commonly asked was "What is your ethnic group/tribe?". Respondents could choose from a number of predefined answer categories, or a category "Other" followed by "Specify". Groups with fewer than 30 observations within a country in a specific year were excluded. Appendix A (table A2.1) provides additional information about the sample used in this study.

We included all fertile women (aged 15-49) with a valid response to the variable preferred family size and for whom the ethnic group was known. Cases with missing data on ethnicity and/or preferred family size, and some unrealistic cases (together 6,7%), were removed from the dataset. Women who gave a non-numeric response (5.4%) such as "It is up to my husband" or "It is up to God", or who did not know how many children they would prefer (8.0%) were excluded. Women preferring more than 12 children (1.2%) were also removed, as we believe that it is unusual for women to give birth to more than 12 children (although it is physically possible).

Missing values on women's education, household wealth and partner's education were addressed by the dummy variable adjustment procedure (Allison, 2009; Little & Rubin, 2019). Missing characteristics on the categorical variables were addressed using an additional "missing" category. The resulting dataset contains information on 499,339 women from 181 ethnic groups living in 24 sub-Saharan countries. Figure 2.2 shows which countries and the number of ethnic groups per country are included.

To get a representative sample of the countries, we used the women weights provided by DHS. We denormalized the women weights and subsequently normalized them according to the national population size of women aged 15-49 for each survey following the standards of ICF International (2012). In this way, a representative sample of the female population aged 15-49 in the 24 countries was created.

Figure 2.2 Sample overview of included countries (N=24) and number of ethnic groups (N=181) for each country.



2.3.2 Methods

To analyze the data, multilevel Poisson regression was used with ethnic group-specific random effects. Multilevel Poisson regression is convenient for count variables, and it provides the opportunity to address the nesting of women within ethnic groups and study the role of factors at the ethnic group level. Fixed effects country dummies were included to control for clustering and confounding at the national level, such as for differences in infrastructure and climate. We use the "variance partition coefficient" (VPC) to describe the share of individuals' total variation in family size preferences that is attributable to the ethnic group level. We calculate the "median rate ratio" (MRR) as described in Austin, Stryhn, Leckie and Merlo (2018) to show the effect size of clustering at ethnic group level.

Four models are estimated. Model 0 shows the clustering at ethnic group level (VPC) without controlling for compositional differences between groups. In Model 1, the group composition is considered by adding characteristics of women and their households. In Model 2, the ethnic group characteristics are added. The final model, Model 3, presents the results of the interaction analysis in which gender-related cultural characteristics of the ethnic group and the group's level of education are interacted with women's individual level of education, position within the household and occupation. Furthermore, the ethnic group's characteristics are interacted with the calendar year.

2.3.4 Variables

The dependent variable "preferred family size" is a count variable indicating how many children women wish to have. In the surveys women were asked: "*If you could go back to the time you did not have any children and could choose exactly the number of children to have in your whole life. how many would that be?".* To ensure that our results were not driven by the social desirability of a match between reached parity and family size preferences, we conducted a robustness analysis in which we reran our analysis based on women aged 15-25 (Appendix table A2.5). Women of this age are less likely to have already reached their full parity, so they are less likely to give a preference based on their reached parity. The results are similar regarding direction and significance, which makes it unlikely that preference was simply based on reached parity.

The independent variables consist of variables measured at household level and ethnic group level. At household level, we control for women's education, their partner's education, women's marital status, women's bargaining power in the household, women's occupation, their partners' occupation, household wealth, place of residence, age, religion, and whether polygamy is practiced.

The education of the woman and of her partner (if available) is measured in years of education completed. A woman's marital status is measured using 4 categories: (1) married or living together, (2) widowed, (3) divorced or not living together, and (4) never married. To indicate a woman's bargaining power in the household, we use the age difference between husband and wife, following earlier studies (Schrijner & Smits, 2018; Spierings et al., 2010). We used 3 categories: (1) equally aged, (2) wife is between 4 and 9 years younger, (3) wife is more than 10 years younger. Women's occupation is measured via a categorical variable indicating whether the woman worked last week (1) in a farming occupation, (2) in a lower non-farming occupation, (3) in an higher non-farming occupation, or (4) did not have an occupation aside

from her housework. Occupation of the partner is measured using the same categories. Household wealth is measured by using the International Wealth Index (IWI; (Smits & Steendijk, 2015), which is a comparative asset-based wealth index. IWI indicates to what extent households own a basic set of assets valued highly by people across the world (like a TV, refrigerator, phone, or car), have access to basic services (water and electricity) and live in a good-quality home. The place of residence is either (0) urban or (1) rural. Religion is categorized as (1) Catholic, (2) Protestant, (3) Muslim, (4) traditional, (5) Christian not specified, (6) no religion and atheist, and (7) other. Women's age is measured by an interval variable. Whether women are in a polygamous relationship is measured with a dummy variable (0=no, 1=yes). To measure the effect of time, we included the calendar year of the surveys in our models.

At the ethnic group level, we include gender-related cultural characteristics and socioeconomic and demographic characteristics. Given that for SSA countries hardly any indicators are available at the ethnic group level, we follow earlier studies that created context variables by aggregating women and household level variables (e.g. Kravdal, 2006; Schrijner & Smits, 2018). We calculated the aggregated mean over all ethnic group members using the household and women weights to create representative indicators for each group. The gender-related cultural characteristics of ethnic groups are measured by the mean age difference between spouses (husband-wife) within the group, by the mean educational differences (men-women) between group members, and by the share of polygamous households. The educational level of the ethnic groups is measured by the mean years of education of group members aged 20-39. The wealth level of groups is indicated by the mean of the International Wealth Index (IWI) of their households. Given that this index at the national level is highly correlated with GDP per capita and the Human Development Index (Smits & Steendijk, 2015), it is expected to be a valid indicator at the ethnic group level as well. The relative size of ethnic groups is measured as their percentage of a country's total population. This variable is meant to control for the possibility that fertility preferences are influenced by relative group size (compare Goldscheider & Uhlenberg, 1969).

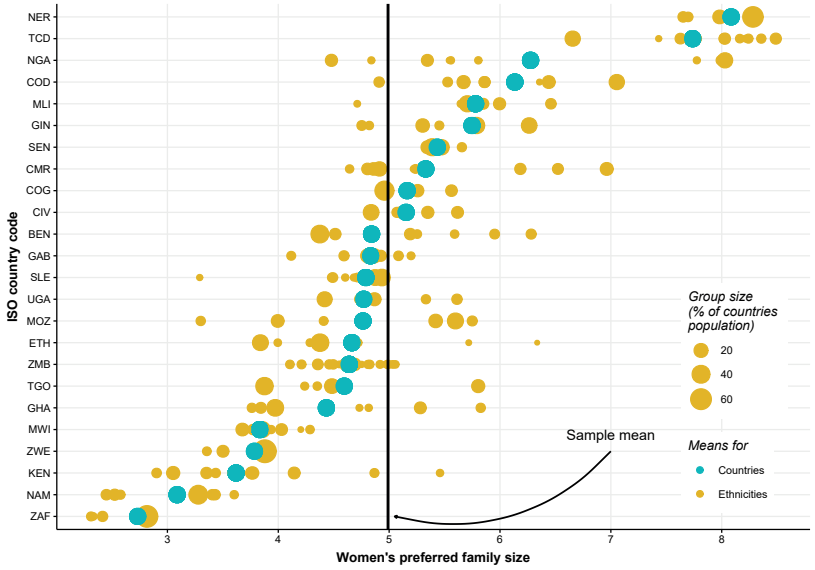
2.4 Results

2.4.1. Descriptive statistics

Figure 2.3 shows the mean preferred family size for each ethnic group (blue dots) and country (yellow dots). The mean number for women's preferred family size

in our sample is 4.99 children. Women in Niger and Chad prefer on average the most children (a mean preferred family size of around 8) and women in South Africa the fewest (a mean preferred family size of below 3). We also observe that the mean preferred family sizes vary substantially between ethnic groups within the countries.

Figure 2.3 Mean preferred family size for ethnic groups, countries and the total sample



Notes: Mean preferred family size (x-axis) for ethnic groups (yellow dots), countries (blue dots) and the total population (vertical black line). The y-axis denotes the ISO code of the different countries. The size of the dots denotes the group size within the countries as a percentage of the countries' population. N ethnic groups=181, N countries=24, N sample=499,339.

Table 2.1 shows the descriptive statistics of all the variables included in our models. The mean number of years' education for women is 5.2, while it is 5.4 for their partners. Most women are married or living together (64%), 28% never married, 3% are widowed and 6% are divorced or no longer living together. Of the married women, most women are between 4 and 9 years younger than their partner (40%), followed by 35% more than 10 years younger. Most women are not working (38%), followed by 32% with a lower non-farming occupation, while most men have a lower non-farming occupation (39%). The average wealth level for each household is 36 on a scale ranging from 0 to 100, and most households (60%) live in rural areas. Women's mean age is 28.1. Most women are Muslim (30%) followed by Protestant (22%). Almost 10% of women live in a polygamous household.

Table 2.1 Descriptive statistics: Means and percentages of variables included in this study

Variables	Mean, %	SD	% Missing
Preferred family size (dep. var.)	4.99	2.57	
Women and household characteristics			
Education woman	5.18	4.56	0.39
Education partner	5.38	5.03	
Marital status			0.17
Married, living together	63.61%		
Widowed	2.58%		
Divorced, not living together	5.79%		
Never married	27.85%		
Age difference husband-wife			2.73
Wife equally aged	21.98%		
Wife between 4 and 9 years younger	40.08%		
Wife >10 years younger	35.21%		
Occupation woman			2.17
Farming	23.94%		
Lower non-farming	32.14%		
Higher non-farming	3.46%		
Not working	38.29%		
Occupation partner			2.38
Farming	41.12%		
Lower non-farming	39.41%		
Higher non-farming	9.39%		
Not working	7.70%		
Wealth of household	36.15	25.65	0.05
Place of residence is rural	59.72%		
Age	28.14	9.29	
Religion			6.45
Catholic	16.86%		
Protestant	22.21%		
Muslim	30.39%		
Traditional	2.57%		
Christian, not specified	18.00%		
No religion, atheist	2.27%		
Other	1.25%		
Woman lives in polygamous household	9.82%		0.00

Table 2.1 Continued

Variables	Mean, %	SD	% Missing
Ethnic group characteristics			
Years of education	5.29	2.82	
Education differences (men-women)	1.62	1.03	
Age differences (husband-wife)	7.80	2.18	
Percentage polygamous households	6.08	0.08	
Average wealth level	33.78	14.88	
Share of population	21.98	17.95	
Time			
Year	2011	4.84	

Source: 1998-2018 DHS (www.dhsprogram.com). N women=499,339, N ethnic groups=181, N countries=24

The descriptive statistics at ethnic group level show that the average number of years of education of group members is 5.3. The mean difference in years of education between men and women in the group (men – women) is 1.6 years and the mean age difference between spouses (husband – wife) is 7.8 years. The mean percentage of polygamous households is 6.1%. The mean wealth level (IWI) of the groups is 33.8. The smallest ethnic group comprises 0.3% of a country's total population and the largest 83.8%. The average ethnic group size comprises 22.0% of the country's population.

2.4.2 Regression results

Model 0 (Table 2.2) tests whether fertility preferences vary between ethnic groups. We observe a significant clustering of women's preferred family size at the ethnic group level. The VPC indicates that 7.4% of the variance in women's preferred family size is due to systematic differences between the ethnic groups. The MRR for this model is 1.12, indicating that women coming from an ethnic group with relatively large family preferences desire at the median 12% more children than women coming from an ethnic group with relatively small family preferences. As such, the results are consistent with our expectations that family size preferences differ between ethnic groups.

Table 2.2 Standardized coefficients of Multilevel Poisson regression analyses on women's preferred family size

	Model 0			Model 1			Model 2		
Variables	*B*	*SE*	**RR**	*B*	*SE*	*RR*	*B*	*SE*	*RR*
Intercept	1.657 ***	0.044	5.242	1.551 ***	0.031	4.717	1.517 ***	0.033	4.559
Year	-0.012 ***	0.001	0.988	0.004 *	0.002	1.004	0.004 *	0.002	1.004
Woman and household characteristics									
Education woman				-0.061 ***	0.001	0.941	-0.062 ***	0.001	0.940
Education partner				-0.005 ***	0.001	0.995	-0.004 ***	0.001	0.996
Marital status, ref.= Married									
Widowed				-0.049 ***	0.004	0.953	-0.045 ***	0.004	0.956
Divorced, not living together				-0.092 ***	0.003	0.912	-0.088 ***	0.003	0.916
Never married				-0.092 ***	0.003	0.912	-0.087 ***	0.003	0.916
Age difference (husband-wife), ref.=equally aged									
Wife between 4 and 9 years younger				0.017 ***	0.002	1.017	0.016 ***	0.002	1.016
Wife >10 years younger				0.016 ***	0.002	1.016	0.015 ***	0.002	1.015
Occupation woman, ref.=farming									
lower non-farming				-0.034 ***	0.002	0.966	-0.035 ***	0.002	0.966
higher non-farming				-0.053 ***	0.004	0.948	-0.054 ***	0.004	0.947
not working				-0.031 ***	0.002	0.970	-0.031 ***	0.002	0.970
Occupation partner, ref.=farming									
lower non-farming				-0.012 ***	0.002	0.988	-0.015 ***	0.002	0.985
higher non-farming				-0.016 ***	0.003	0.984	-0.019 ***	0.003	0.981
not working				-0.017 ***	0.005	0.983	-0.022 ***	0.005	0.978
Wealth of household				-0.038 ***	0.001	0.962	-0.037 ***	0.001	0.963
Place of residence, ref.=urban				0.048 ***	0.002	1.050	0.049 ***	0.002	1.050
Age				0.078 ***	0.001	1.081	0.079 ***	0.001	1.082

Table 2.2 Continued

	Model 0			Model 1			Model 2		
Variables	*B*	*SE*	**RR**	*B*	*SE*	*RR*	*B*	*SE*	*RR*
Religion, ref.=Catholic									
Protestant				0.027 ***	0.003	1.028	0.027 ***	0.003	1.028
Muslim				0.083 ***	0.003	1.087	0.083 ***	0.003	1.086
Traditional				0.091 ***	0.005	1.096	0.092 ***	0.005	1.096
Christian, not specified				0.002	0.002	1.002	0.002	0.002	1.002
No religion, atheist				0.049 ***	0.006	1.050	0.050 ***	0.006	1.051
Other				0.048 ***	0.008	1.049	0.047 ***	0.008	1.048
Woman lives in polygamous household				0.023 ***	0.002	1.023	0.025 ***	0.002	1.025
Ethnic group characteristics									
Years of education							-0.029 ***	0.007	0.971
Education differences (men-women)							0.033 ***	0.003	1.034
Age differences (husband-wife)							-0.025 ***	0.005	0.975
Percentage polygamous households							-0.004	0.004	0.996
Average wealth level							0.012	0.006	1.012
Share of population							0.029 ***	0.004	1.029
Variance of random intercept		0.015			0.007			0.007	
Deviance		2,091,315			2,030,763			2,030,627	
VPC		0.074			0.033			0.033	
MRR		1.124			1.085			1.086	

Notes: * $p<.05$, ** $p<.01$, *** $p<.001$; N women=499,339, N ethnic groups=181. The continuous variables are standardized. Both models include the full set of country-level fixed effects dummies to control for clustering at the national level (Appendix A, table 2.3).

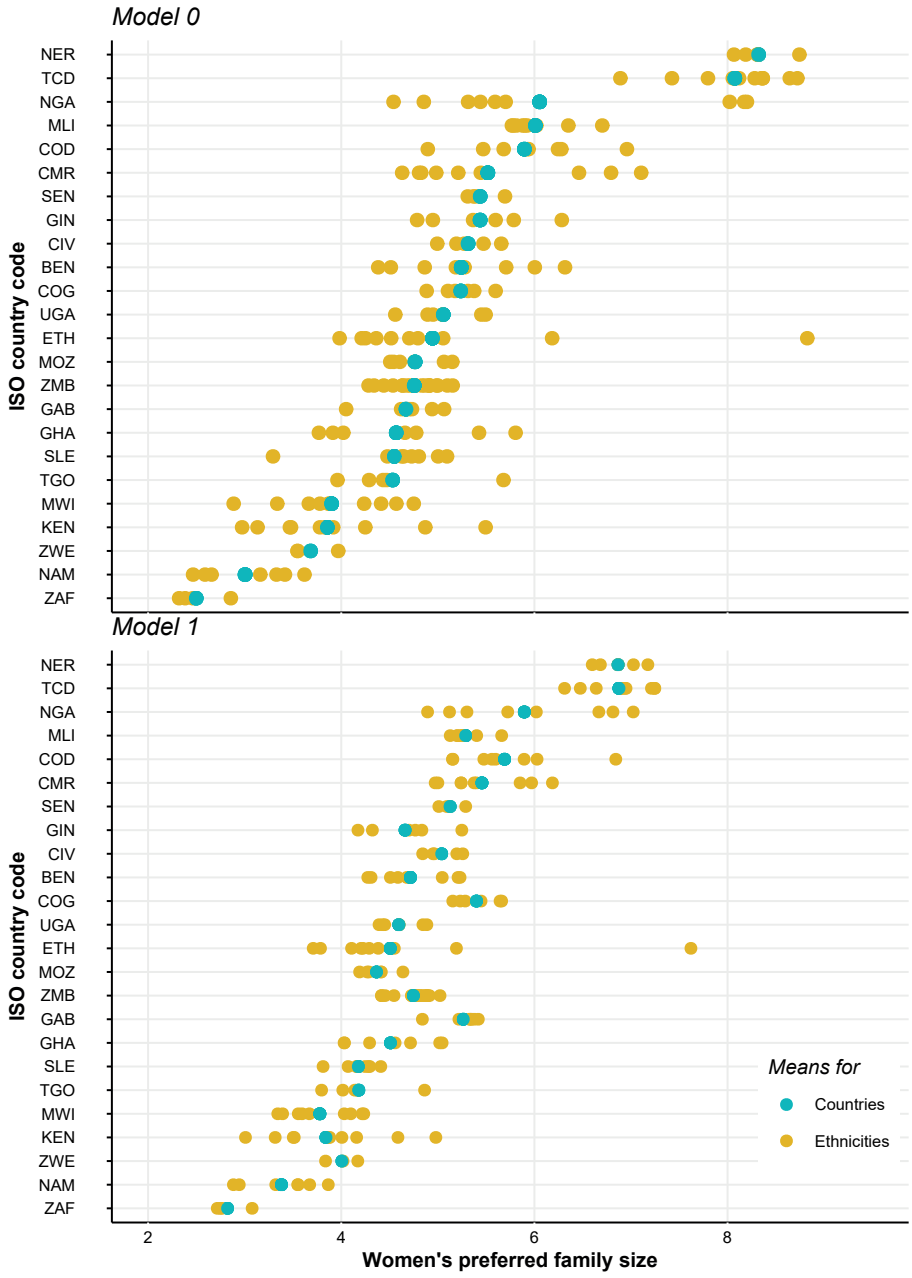
Model 1 (Table 2.2) tests whether the variation between ethnic groups is a consequence of variation in socioeconomic, demographic, and cultural characteristics at the individual and household level, as proposed by the social characteristics' hypothesis. Both the Poisson regression coefficients (*B*) and the rate ratios (RR) are presented. The RR can be most easily interpreted. For example, a RR of 0.94 for women's education means that if women's education increases by 1 standard deviation, the preferred number of children decreases by a factor of 0.94.

The coefficient of the women's and household characteristics are largely in line with previous studies (Atake & Gnakou Ali, 2019; Bongaarts, 2020; Kebede et al., 2019; Lerch, 2019; Matovu et al., 2017; Tomkinson, 2019). Women prefer fewer children when they or their partners have a higher level of education, when they are not married, are the same age as their partner, when they or their partners have a non-farming occupation, their households are wealthier, they live in urban areas, they are younger, they are Catholic, and when they are not living in a polygamous relationship.

A surprising finding is that the year coefficient is positive, suggesting that women have come to prefer more children in recent years. The bivariate relationship between time and family size preferences is negative (see Model 0). The coefficient becomes positive only after adjusting for other factors. Further explorations showed that this happened when women's education and household wealth were included. The positive year coefficient is thus likely to be the result of an increase in women's education and wealth over time.

After inclusion of women's and household characteristics, the VPC decreased from 7.4% to 3.3%. Visual comparison of mean predicted preferred family sizes for ethnic groups of Model 1 with Model 0 (figure 2.4) shows a decrease in variance of ethnic group means around the country means. As such, the variance between ethnic groups from model 0 is partly explained by compositional differences between ethnic groups, but a substantial difference in family size preferences between ethnic groups remains. The MRR decreased to from 1.12 to 1.09. Hence, women from a group with relatively large family preferences desire 9% more children at the median than women from a group with relatively small family preferences when compositional differences are taken into account.

Figure 2.4 Mean preferred family size for ethnic groups, countries and the total sample with control variables



Notes: Mean preferred family size (x-axis) for ethnic groups (yellow dots) and countries (blue dots). The y-axis denotes the ISO code of the different countries. The upper graph shows the predicted values from model 0, and the lower graph shows the predicted values from model 1 where the ethnic groups show less variance around the country means. N ethnic groups=181, N countries=24.

2.4.3 Ethnic group characteristics

Model 2 adds ethnic group characteristics to test to what extent these predict women's preferred family size. The inclusion of ethnic group characteristics is a significant improvement on the previous model. The deviance of Model 2 versus Model 1 is significant smaller ($X^2(6,N)=170,0$, $p<.001$). As such, ethnic group characteristics are associated with women's preferred family size.

The coefficients of the group characteristics (Model 2, Table 2.1) show that women prefer fewer children when they are members of a group with an average higher level of education. This finding is in line with our expectation that, independent of one's own level of education, the group's level of education has an additional influence. Women prefer larger families when they live in groups with larger education differences (men – women) at the ethnic group level, while women prefer smaller families when they live in groups with larger age differences between spouses (husband – wife). Women's preferred family size is not influenced by the percentage of polygamous households at the group level, so polygamy seems only to influence women's preferences at the household level. These findings show mixed evidence for our hypotheses, as we expected these gender-related cultural characteristics to be positively related to family size preferences. The average level of wealth is not significantly related to family size preferences, but the relative size of the group within the country is positively related to family size preferences. Women prefer more children when they are members of a group that is relatively large compared to other groups within the country.

2.4.4 Moderations

Model 3 (Table 2.3) shows the results of our interaction analysis in which we test whether the effects of socioeconomic and demographic characteristics of women at the household level are moderated by the level of education and gender-related cultural characteristics at the group level.

Table 2.3 Standardized interaction coefficients between major group and individual characteristics

	Model 3		
	B	*SE*	*RR*
Education differences (men-women) at group level			
** Individual characteristics*			
level of education woman	0.010 ***	0.001	1.010
occupation woman: lower non-farm	-0.007 **	0.002	0.993
occupation woman: higher non-farm	-0.028 ***	0.005	0.973
occupation woman: not working	-0.012 ***	0.002	0.988
age difference: wife between 4 and 9 years younger	-0.005 *	0.002	0.995
age difference: wife >10 years younger	-0.007 ***	0.002	0.993
year	0.003 **	0.001	1.003
Age differences (husband-wife) at group level			
** Individual characteristics*			
level of education woman	0.009 ***	0.001	1.009
occupation woman: lower non-farm	-0.009 ***	0.003	0.991
occupation woman: higher non-farm	-0.040 ***	0.005	0.960
occupation woman: not working	-0.007 **	0.002	0.993
age difference: wife between 4 and 9 years younger	-0.023 ***	0.002	0.977
age difference: wife >10 years younger	-0.019 ***	0.002	0.981
Year	0.010 ***	0.001	1.010
Years of education at group level			
** Individual characteristics*			
level of education woman	0.000	0.001	1.000
occupation woman: lower non-farm	-0.014 ***	0.002	0.986
occupation woman: higher non-farm	-0.033 ***	0.006	0.968
occupation woman: not working	-0.008 ***	0.002	0.992
age difference: wife between 4 and 9 years younger	-0.020 ***	0.002	0.980
age difference: wife >10 years younger	-0.016 ***	0.002	0.984
Year	0.000	0.001	1.000
Variance of random intercept		0.007	
Deviance		2,029,714	
VPC		0.031	
MRR		1.083	

Notes: * $p<.05$, ** $p<.01$, *** $p<.001$; N women=499,339, N ethnic groups=181. The continuous variables are standardized. The model includes the full set of women- and household characteristics, ethnic group characteristics and country-level fixed effects dummies (presented in Appendix A, table 2.4).

We find that the negative effect of women's individual education on their preferred family size is stronger in groups where the gender differences in education and the age differences between husbands and wives are smaller. This means that the effect of education is reduced for women living with greater gender differences, as predicted by the situational dominance hypothesis. The effects of having a non-farming occupation are also significantly related to ethnic group characteristics, but in the opposite direction: they are stronger if the differences are greater. This shows that women's occupation is especially important where there are larger gender differences, which is in line with the individual dominance hypothesis. We further find that the effect of age difference between spouses at the household level is smaller in groups with larger age differences. The positive effect of age difference at household level is therefore reduced where there are greater gender differences, as proposed by the situational dominance hypothesis.

The effect of women's education on family size preferences is not moderated by the years of education at group level, but women's occupation and households' age differences is. The negative effects of having a non-farming occupation are stronger in groups that are on average more highly educated, which supports the situational hypothesis. The positive effects of household age differences are weaker if groups are on average more highly educated, which is in line with the individual dominance hypothesis. Finally, we see that the positive effect of education difference has become greater in more recent years. The negative effect of age difference at ethnic group level has lessened in recent years, whereas the effect of years of education at the group level has remained stable.

2.5 Discussion

The present study aimed to ascertain to what extent women's family size preferences are influenced by ethnicity and what specific characteristics are important. Based on 500,000 women aged 15-49, belonging to 181 ethnic groups within 24 sub-Saharan countries, we found broad evidence that ethnicity is associated with women's preferred family size. Our analyses revealed that the influence of ethnicity almost halved (from 7 to 3%) when individual characteristics were taken into account, but that there remained a substantial structural influence. Ethnic group characteristics that were found to be important were gender-related cultural characteristics and the mean level of education. The size of these effects were found to be moderated by women's individual characteristics. We therefore conclude that ethnicity has a substantial influence on women's preferred family size.

Regarding the way in which fertility outcomes are influenced by ethnicity, two major hypotheses were tested: the s*ocial characteristics hypothesis*, which assumes ethnic fertility differentials to be solely due to compositional differences between groups (Addai, 1999; Bakibinga et al., 2016; Petersen, 1969) and the *cultural hypothesis*, which assumes that besides compositional differences, norms and values at the ethnic group level are also important for fertility outcomes (Caldwell & Caldwell, 1987; Kollehlon, 1989; Palamuleni, 2014). Our findings, which showed a substantial association between ethnic group characteristics and family size preference, are in line with the cultural hypothesis. This adds to the growing body of evidence indicating that ethnicity influences women's fertility behavior through group pressure and social support (Adedini et al., 2015; Berndt et al., 2018; Levy & Nail, 1993; Lois & Arránz Becker, 2014).

Results for our expectations regarding the direction of the effect of gender-related cultural characteristics were mixed. While we expected that larger educational differences, larger age differences and a greater number of polygamous relationships would relate to larger family size preferences, we only found this effect for educational differences. We found the opposite effect for larger age differences, and we did not find an effect in relation to polygamous relationships. Although surprising, this finding is in line with earlier studies that have reported inconsistent effects of gender-related cultural characteristics on fertility behavior (Moursund & Kravdal, 2003; Upadhyay et al., 2014; Upadhyay & Karasek, 2012).

Our interaction analysis revealed that the influence of women's individual characteristics depended to a certain extent on characteristics of their ethnic group. The negative relation between women's education and larger families is weaker in ethnic groups with greater gender differences in terms of education and age difference between spouses. This outcome is in line with the situational dominance hypothesis, which predicts that individual characteristics can be less important in groups with greater gender differences. This is in line with earlier findings of Spierings et al. (2010) which examined the potential for women to be gainfully employed in patriarchal countries. At the same time, we found evidence for the individual dominance hypothesis, as our results showed that the effect of a non-farming occupation and women's bargaining power in the household is stronger in groups with greater gender differences.

Given our findings, we recommend that policy-makers in this field examine their policy target and look not only at women's individual characteristics but also at the circumstances in which they are living, including those of the ethnic group

to which they belong. For example, our finding that women's education has a larger influence in groups where gender-related cultural differences are smaller indicates that policies need to incorporate the social context as well. More detailed information on interactions between individual characteristics and the social context will be especially beneficial, as this would inform policy-makers as to which individual characteristics are important to focus on in specific contexts.

Some caution is required regarding our conclusions. First, with respect to the data, some ethnic groups did not have sufficient respondents in the DHS surveys and as such, could not be included in our analysis. Nevertheless, our study is a major improvement on previous studies, as we included many more ethnic groups than others have done so far. Second, we must bear in mind that the strength of ethnic identity might vary between individuals and that ethnicity is not the only characteristic with which women identify. Characteristics such as social class, religion, and national identity are also likely to be important. Third, we should be aware that a woman's ideal family size is often a response to the dynamics of her current situation, such as change in marital status (divorce, widowhood, and new marriage) (e.g. Sennott & Yeatman, 2012; Trinitapoli & Yeatman, 2018). Moreover, the lives of African women are often unpredictable, which could be reflected in dynamic fertility preferences over time (Johnson-Hanks, 2005; J. A. Johnson-Hanks et al., 2011). As such, fertility preferences should not be seen as the basis for a feasible plan for the future but rather as a reflection of that specific moment, and care should be taken not to draw conclusions regarding fertility rates from these preferences.

In sum, this study provides evidence that ethnic groups have a substantial influence on women's family size preferences, even after controlling for group composition. Compared with earlier studies, our research constitutes a major improvement because it provides – for the first time – a broad comparative analysis of the variation in preferred family size between ethnic groups. The findings clearly indicate that ethnic background should not be forgotten when designing policies that aim to affect fertility (preferences).

Chapter 3: The Role of Norm Strength and Social Pressure for Decision-Making: Evidence on the timing of marriage and childbirth in sub-Saharan Africa[2]

[2.] This paper is based on joint work with Natascha Wagner. This chapter was presented at the Culture and Economics Workshop in Nijmegen 2023.

Early marriage and adolescent childbearing remain widespread in Sub-Saharan Africa (SSA), despite strong evidence of their harmful consequences and extensive policy efforts to curtail the practices. This persistence is in part driven by social pressure to conform to traditional family formation norms. This paper introduces a novel proxy to measure social pressure using the coefficient of variation (CV) in the timing of marriage and first birth within a woman's social groups – defined by neighborhood, ethnicity, and religion. Drawing on 104 geo-referenced Demographic and Health Surveys (1994–2023) from 33 SSA countries, we show that greater variation in the timing of family formation is associated with delayed transitions and reduced risk of early marriage and teen pregnancy. A one standard deviation increase in group-level variation corresponds to a 2-3 months delay in marriage and childbirth and a 1.8–1.9 percentage point reduction in the likelihood of early transitions. Variation in older cohorts predicts the timing of marriage, but not childbirth, suggesting that inter-generational pressure is strong for marriage. In turn, norms around fertility are shaped more by peers. Effects are strongest among rural and less-educated women, suggesting education and urban living can mitigate normative pressure. Our approach offers a readily scalable way to capture local norm strength and helps explain why behavior change remains limited in some contexts despite improvements in overall economic conditions, exposure to other lifestyles and advances in legal protection.

3.1 Introduction

Cultural traditions play a critical role in shaping family formation behavior in sub-Saharan Africa (SSA), particularly by influencing practices related to marriage and childbirth (Bertocchi et al., 2024). Researchers are increasingly recognizing the persistence and continued impact of these norms, even when individuals and communities are faced with evidence of their negative consequences (Akerlof & Kranton, 2000; Henrich, 2015). Norms in general serve as identity markers and foster social belonging and especially norms around sexual transitions and family formation tend to be influential (Akerlof & Kranton, 2000). For example, SSA has the highest rates of child marriage and teen pregnancies worldwide, showing that within communities that practice the tradition identity and reputation mechanisms are at play (Akerlof, 1980). However, both traditions are associated with serious health and developmental risks for adolescent girls and their children (Chari et al., 2017; Field & Ambrus, 2008; Ganchimeg et al., 2014; Jensen & Thornton, 2003; WHO, 2024). In addition, the persistence of these harmful practices undermines global development goals, especially the Sustainable Development Goal (SDG) 5, which aims to eliminate child marriage and teen pregnancy. Despite widespread awareness of the negative implications, policies aimed at reducing child marriage and teen pregnancy have often been ineffective (Feyissa et al., 2023).

As forcefully demonstrated for the case of other harmful practices - female genital mutilation - one possible explanation for the limited effectiveness of empowerment policies targeting adolescent girls in SSA is the influential role of the community and social pressure to comply with existing traditions (Bellemare et al., 2015; Coyne & Coyne, 2014; Wagner, 2015). Social pressure in the context of harmful traditional practices refers to both informal and formal mechanisms by which individuals are influenced - consciously and unconsciously - to adhere to culturally ingrained expectations regarding family formation (Asadullah & Wahhaj, 2019; Tewahido et al., 2022). These pressures may arise from family, peers, religious groups, and the broader community (Bicchieri et al., 2014; Wilkinson et al., 2024).

While social pressure has been repeatedly identified as the key factor behind the persistence of culturally harmful norms, a major challenge is quantifying social pressure empirically resulting in limited empirical evidence on the causal role of norms. To address this gap, we propose a novel approach that examines variations in the age of marriage and childbirth within communities. We hypothesize that greater variation in their timing reflects weaker social norms, implying less pressure to conform and greater individual autonomy in family formation choices.

Our novel, readily scalable measure of social pressure is grounded in observed behavioral variation. The core premise is that variation in the timing of marriage and first birth within social groups serves as a proxy for the strength of normative expectations: lower variation reflects stronger norms and higher social pressure, while greater variation signals weaker enforcement of norms and increased individual agency. We operationalize this concept by calculating the coefficient of variation (CV) in family formation behavior within three social contexts – defined by the neighborhood community, ethnicity, and religion – and examine how this variation relates to women's family formation trajectories. Our analysis draws on marriage and birth data, focusing on two samples of 224,806 and 548,428 women from 33 countries and 104 DHS surveys, with data spanning from 1994 to 2023. For each woman, we calculate the CV of her social surroundings based on neighborhood, ethnicity, and religion. We then test whether the variation in the timing of family formation as represented by the CV can explain her timing of marriage and childbirth, including the chances of becoming a child bride or a teen mother using linear probability models. Contrary to self-reported measures of social norm perception our measure has the advantage of being free from social desirability bias and thus providing an objective measure of the actual extent of the practice of child marriage and teen pregnancy.

Our results consistently show that 'normative variation' in a woman's social surroundings serves as an explanation for a woman's timing of marriage and childbirth. Across all our models, we find that women living in neighborhoods, religious and ethnic groups with more variation in the timing of marriage and childbirth tend to start their family formation later. Specifically, women living in diverse neighborhoods, ethnic, and religious groups (measured as one standard deviation above the mean) marry, on average, 3 months later, and give birth to their first child 2 months later. Moreover, these women are 1.9 percentage points less likely to become child brides and 1.8 percentage points less likely to become teen mothers. Additionally, we find that not only the current variation in the lived group norm but also the normative variation of the previous generation plays a non-negligible role for women's timing of marriage, whereas this is not the case for women's timing of childbirth. Moreover, we find that for the extreme forms of early family formation, child marriage and teen pregnancy, women's level of education and residence can (partly) overcome the group pressure to engage in such behaviors. The results are robust to employing different neighborhood sizes, the inclusion of household wealth and the exclusion of countries from the analysis, which the latter suggesting that it is not one single country that drives the results.

We contribute to a growing body of literature on the role of social norms in shaping marriage and fertility decisions. Previous studies have shown that while gender and marriage norms are evolving in some contexts, these changes are less pronounced in highly gender-stratified societies (Batyra et al., 2021). Similarly, there is evidence that early child marriage (defined as marriage before the age of 15) has remained stable over the past decade, while late child marriage has declined, indicating a strong social pressure to conform to the norm of early child marriage in certain cultures (Koski et al., 2017). For the case of Niger, for example, Sripadetal. (2024) show that norms around the protection of women from harassment, the fear of socially induced ruined marital prospects, and the neighbors' take are non-negligible factors in the decision-making about early marriage. In this regard, differences in cultural norms may also explain the varying impact of economic changes, such as increases in wealth and education, on marriage and the timing of childbirth. For example, CastroTorresetal. (2022) identified that increases in the age at first birth were unevenly distributed between regions, potentially due to different cultural norms on family formation. In addition, cultural norms have been identified as significant determinants of child marriage in certain contexts, such as Tanzania and South Africa (Ahonsi et al., 2019; Schaffnit et al., 2019). Building on this literature, our study advances the understanding of how social pressure – both to marry early and to conform to cultural norms around childbirth – affects the timing of marriage and fertility decisions.

More specifically, our study contributes to the measurement of social pressure - or the strength of adherence to social norms - within social groups. While qualitative studies often attempt to measure cultural traditions directly, they are not suitable for large-scale comparative research. Quantitative studies have sought to measure the strength of cultural norms using the cultural tightness-looseness indicator, which operates on a national scale (Fischer, 2009; Gelfand et al., 2006). This indicator is defined by the degree of adherence to norms on the one hand, and the severity of punishments for norm violations on the other (Gelfand et al., 2011). However, this indicator is better suited to Western societies and is not effective for measuring cultural strength within specific sub-groups, such as ethnic and religious groups. This poses a challenge for understanding family formation trajectories in SSA, where sub-groups play a critical role in shaping social norms (Ahonsi et al., 2019; Preston-Whyte & Zondi, 1991; Stevanovic-Fenn et al., 2015; Zgambo, 2022). By focusing on dynamics within these smaller groups, our study develops a more localized measure of social pressure, enabling a finer-grained understanding of cultural norms across different social groups while being applicable and comparable across different contexts and for large samples.

The remainder of the paper is structured as follows. Section 3.2 provides background information on marriage and childbirth in SSA. We also elaborate on variations within social groups as a proxy for social pressure to adhere to social norms. Section 3.3 introduces the theoretical model that motivates the study, section 3.4 presents the data and construction of the estimation sample. The estimation strategy is introduced in section 3.5 and the results are discussed in section 3.6. Section 3.7 concludes.

3.2 Background

3.2.1 Marriage and fertility in SSA

Marriage has long been one of the most respected institutions in African cultures, often marking the beginning of women's reproductive lives. In SSA, marriage has historically been near universal for both men and women, particularly during the 1950s and 1960s. Women, in many communities, were considered eligible for marriage after reaching puberty. Traditional marriage patterns in SSA also involved significant age differences between husbands and wives, and polygyny, where one husband is married to multiple wives, was widely practiced despite its negative impacts on the women and children born into polygynous families (Garenne, 2022; Wagner, 2015).

However, by the late 20th century and starting in the post-independence period in the 1960s, marriage patterns began to shift. Notably, the average age of first marriage for women has increased, while overall marriage rates have declined for both men and women (Shapiro & Gebreselassie, 2014). The rising levels of women's education and income are recognized as key drivers of these changes, particularly influencing the delay in marriage (Garenne, 2004). Despite these shifts, SSA still has the highest rates of child marriage in the world. Nine out of ten countries with the highest child marriage rates are located in SSA (UNICEF, 2015). Early marriage remains a significant barrier to development, with well-documented negative consequences for women's education, maternal and child health, intimate partner violence, and women's labor market participation (e.g. Chari et al., 2017; Field & Ambrus, 2008; Hicks & Hicks, 2019; Jensen & Thornton, 2003).

Alongside marriage, childbearing is a central aspect of family formation in SSA. In many African societies, motherhood is not only a highly valued role, but also seen as a key component of a woman's identity and social status (Ahonsi et al., 2019; Schaffnit et al., 2019). Women in SSA tend to give birth to their first child at an earlier age compared to women in other low- and middle-income countries (Kassa et al., 2018; UNICEF, 2015). As a result, SSA has the highest rates of adolescent pregnancy

worldwide (Kassa et al., 2018; WHO, 2024; Yakubu & Salisu, 2018). For example, in 2023, the adolescent birth rate for SSA girls aged 10-14 years was 4.4 per 1,000 women, and the rate for girls aged 15-19 years was 97.9 per 1,000 women, significantly higher than the global average of 1.5 and 41.3 per 1,000 women, respectively.

Teen pregnancy in SSA is associated with various adverse health outcomes for both mothers and children, including low birth weight, premature birth, anemia, pregnancy-induced hypertension, and increased infant mortality. Moreover, akin to early marriage adolescent pregnancy is often linked to school dropout, poverty, and increased vulnerability to domestic violence (Branson & Byker, 2018; UNICEF, 2015).

Despite the negative consequences of early marriage and childbearing, the practices are widely upheld in SSA. Their persistence in the region reflects broader societal expectations and pressures that promote early marriage and childbearing, highlighting not only the important role of social norms and resulting pressures and expectations, but also the interconnectedness of marriage and fertility. This is why we study both phenomena jointly.

3.2.2 The role of social norms and social pressure

Social norms within groups, such as ethnic and religious communities, play a major role in shaping daily life in SSA. Until today, coming of age and key life events such as marriage and family formation are not only traditionally loaded, but also marked with traditional symbols, attributes, rituals, and rites of passage (Mbaye & Wagner, 2017; Ofori & Mohangi, 2024; Schroeder et al., 2022). Accordingly, women's decisions regarding marriage and childbearing are often influenced, and in many cases constrained, by expectations from family, peers, and the broader community (UNICEF, 2015; Varga, 2003; Zgambo, 2022). These communities set the standards for when and how women should marry and give birth, with strong community pressure to conform to these widely accepted practices and timelines. Despite changes in marriage and childbearing patterns in recent decades, these social expectations continue to drive the persistence of early marriage and teen pregnancy in many SSA communities.

A growing body of literature highlights how these expectations are maintained and transmitted through social networks, where non-compliance can result in reputation loss or exclusion. In many cases, the perceived benefits of adhering to tradition outweigh the potential gains from delaying marriage or childbearing. Asadullah and Wahhaj (2019) show how norms around early marriage are reinforced through peer and family networks, while Chari et al. (2017) find that

families may comply with early marriage expectations even when aware of its negative effects on child outcomes. Bellemare et al. (2015) demonstrate that similar mechanisms sustain female genital cutting, where social pressure within communities perpetuates harmful norms. Moreover, regional variation in family formation trajectories is often explained by differences in normative contexts rather than socioeconomic conditions alone (Castro Torres et al., 2022). In SSA, this dynamic is particularly visible. In Niger, both empirical expectations (what others do) and normative expectations (what others think one should do) shape marriage decisions, especially through the influence of elders and religious leaders (Sripad et al., 2024). Likewise, Tewahido et al. (2022) report that in rural Ethiopia, early marriage is widely viewed as a moral obligation, and girls marry to avoid stigma and preserve family honor. Similar motivations are documented in Ghana and South Africa, where early motherhood is seen as a means of attaining respect and fulfilling culturally prescribed roles (Ahonsi et al., 2019; Mkhwanazi, 2010; Preston-Whyte & Zondi, 1991). Together, these studies underscore that early family formation is not simply a function of poverty or limited opportunity, but a socially reinforced response to deeply embedded cultural expectations.

Cultural norms, in turn, vary greatly between social groups, both in terms of their content and the strength of their enforcement. No single religious affiliation universally dictates family formation practices. Child marriage and teen pregnancy are more prevalent in some predominantly Christian, Muslim, Hindu, or Buddhist regions than in others (UNICEF, 2015). In fact, within countries or regions, religion can be a significant factor in explaining the rates of child marriages and teen pregnancies within certain areas, while having little influence in others (ICRW, 2007). Similarly, ethnicity plays a crucial role in determining marriage and childbearing practices. For example, child marriage is common among the Hausa ethnic group in Niger but much less prevalent among the Tuareg (UNICEF, 2015). These disparities highlight the importance of considering social background at a regional level when examining the influence of cultural practices on family formation in SSA.

The strength of cultural norms also affects the degree of conformity in family formation trajectories. In societies with tightly enforced norms, social pressure and the threat of sanctions for deviation are stronger, which often results in more uniform timing of marriage and childbearing. For instance, in rural Tanzania, women who marry ‘too late’ are frequently stigmatized and regarded as failing to meet societal expectations (Schaffnit et al., 2019). Consequently, we hypothesize that women in communities with rigid cultural norms are more likely to follow traditional family formation patterns, leading to less variation in age at marriage

and first birth. In contrast, looser normative environments likely allow greater individual autonomy, which should manifest in more variation in these outcomes. Therefore, we consider observed variation in the timing of family formation within social groups as an indicator of underlying normative flexibility and, by extension, of the strength of social pressure.

3.3 Conceptual model: Social image utility model of early marriage and teen pregnancy

Conceptually, we root ourselves in the social image utility models. Individual actions, for example, the involvement in early marriage or early childbirth, are a function of expected social desirability within a social group, the degree to which an individual cares about the judgment of the groups, and the individual benefits and costs of the behavior (Bursztyn & Jensen, 2017). In this class of models, if a woman decides to marry or have a child, the advantages associated with these actions outweigh those of not engaging in the practices.

We start deriving the model by detailing the concept of social image. Many actions that one might undertake are observable by others and may send signals about one's type. This is especially the case for marriage and fertility. An individual exhibits social image concerns when her utility depends on the expectations about her type that are held by others, conditional on observing her behavior. Formally, consider an individual *i* with reference group *j* undertaking a visible action $a_i \in (0, 1)$ that might reveal information about *i*'s type $\sigma_i \in (l, h)$. Let us assume that being of type *h* (high type) is more socially desirable within the reference group *j* than being of type *l* (low type). We define the social image term S_{ij} in *i*'s utility function as:

$$S_{ij} = \lambda_{ij} E_i(\omega_j) Pr_{i(\sigma_i = h|\alpha_i)} \qquad (3.1)$$

In this equation, $Pr_{i(\sigma_i=h|\alpha_i)}$ represents the probability that other individuals in the reference group think that *i*'s type is *h* conditional on observing *i*'s choice of behavior, a_i. The term $E_i(\omega_j)$ corresponds to *i*'s expectation about the social desirability, ω_j, of being considered by group *j* as type *h*. Being of type *h* is socially desirable if $\omega_j > 0$. Finally, λ_{ij} measures how much individual *i* cares about being perceived by group *j* as being of type *h*. If $\lambda_{ij} > 0$, then *i* is a conformist; if $\lambda_{ij} < 0$, then *i* is a contrarian; and if $\lambda_{ij} = 0$, then *i* does not care about her image or what others think about her type. For any non-zero level of λ_{ij}, individual *i* is subject to concerns about social image and is therefore subject to social pressure from group *j*.

There are two points we would like to emphasize within this framework. First, different groups may have different ω_j's. For example, in our case, some groups may consider child marriage as socially desirable, whereas other groups may have a different take. Second, a woman may care about how she is perceived by more than one reference group. For example, a woman jointly belongs to a religious community, ethnic group, and to her regional community. Thus, her decision making may be influenced by countering forces. For simplicity, we only model the impact of one group's ω_j. This one group's expectation about social desirability ω_j can also be considered as the weighted average across all groups that the individual cares about.

Image concerns are not the only driver of behavior. To consider the empirical implications of social image, we embed the conceptual framework within a random utility model. Assume there is a latent variable $a^{\sim}_i$ capturing the desirability of some behavior a_i. The underlying utility is a function of the social image component, as described above, as well as some benefits, B_i, and costs, C_i, which are unrelated to the image concerns and may be individual specific:

$$\tilde{\alpha}_i = B_i - C_i + \lambda_{ij} E_i(\omega_j) Pr_{i(\sigma_i = h|\alpha_i)} + \varepsilon_i \qquad (3.2)$$

where ϵ_i is a random variable. The researcher observes whether the socially desirable behavior, in our case marriage and childbirth, was lived depending on whether the (unobserved) net utility is positive or negative, so the actual behavior $a_i = 1$ if $a^{\sim}_i \geq 0$ and $a_i = 0$ if $a^{\sim}_i < 0$.

In our specific case, we argue that diversity in behavior is the consequence of a lower social desirability of the observed behavior (ω_j) of being of type *h* in the group and diminishes an individual's belief about the social desirability ($E_i(\omega_j)$) of that behavior as well. As a consequence, the importance of social image in the desirability of a behavior diminishes and the individual benefits (B_i) and costs (C_i) become more important, which will likely result in a more diverse timing of family formation.

Importantly, the social image utility model makes it apparent that adolescent girls or their families may opt for early marriage, or the girls might engage in risky sexual behavior and become pregnant, even if they do not directly benefit from their action. Yet, the utility benefit is derived from the projection of a desired image to others and the avoidance of negative social consequences. Put differently, when

considering the role of social norms for decision-making not only the personal gain and thus the direct utility impact is considered but also the utility derived from projecting a positive social image. This projection effect can vary from society to society and even from social group to social group despite otherwise similar socio-economic background conditions. The impact hinges on the groups social desirability parameter, ω_j. Empirical we conceptualize the social desirability parameter as coefficient of variation in the observed behavior of the different groups - neighborhoods, ethnic, and religious groups - the adolescent girls are part of and test for its impact on realized individual behavior in a linear regression framework. Thus, empirically we address the relevance of the core parameter that conceptualizes the social image model. Next we turn to the data we use for the empirical analysis.

3.4 Data

To address whether there is a role for social pressure in early marriage and teen pregnancy, this study relies on 104 geo-referenced Demographic and Health Surveys (DHS)[3] for 33 SSA countries between 1994 and 2023. DHS are large, nationally representative household surveys[4]. Using geo-referenced DHS surveys has the advantage that clusters can be aggregated to new units of analysis like local neighborhoods, rather than being constrained by the survey to rely on larger administrative units like provinces[5], allowing for the empirical analysis to be more granular. A complete overview of the surveys included in the analysis can be found in Appendix B Table A3.1.

3. www.dhsprogram.com

4. The sampling strategies of these surveys – to create a nationally representative, randomly selected sample of households – are to first randomly select a subset of non-overlapping geographical areas, called 'clusters', which typically comprise communities, villages, towns, cities, or neighborhoods. In each of the selected clusters, all households are listed, and a random subset of 25-30 households is selected to participate in the surveys. Household members are first asked for basic household information, including a household roster before women aged 15 to 49 years are interviewed about their marriage and birth history.

5. DHS applied a geographic displacement procedure to maintain the confidentiality of survey respondents. For urban clusters, the DHS displaced the coordinates of the clusters up to two kilometers, and for rural clusters up to five kilometers, with a further, randomly selected 1% of rural clusters displaced a distance up to 10 kilometers. Overall, the displacement process produces data with uniformly distributed displaced distances (Burgert et al., 2013), making it unlikely that the displacement procedure influences the results.

We deploy four dependent variables: (a) a woman's age at first marriage, (c) whether a woman was a child bride, (c) her age at first birth, and (d) whether she became a mother as a teenager. A woman's age at first marriage and her age at giving birth to the first child are measured in whole years, with a minimum age of 10 years. Whether a woman is a child bride is measured as marrying before the age of 18 (1 = yes; 0 = no), and whether a woman is a teen mom is measured as giving birth before the age of 20 (1 = yes; 0 = no). For deriving these cut-offs, we implemented the formal definitions of the World Health Organization for child marriage and teen pregnancy.

To capture social pressure to conform to prevailing norms around the timing of family formation, we construct measures of variation in the timing of marriage and childbirth within three different social groups: (1) local neighborhoods, (2) ethnic groups within regions, and (3) religious groups within regions. For local neighborhoods, we use the geographic coordinates of DHS clusters to define a circular area with a radius of 50 kilometers around each cluster. All other clusters within this radius constitute a woman's local neighborhood. This flexible spatial approach allows us to capture variation in proximate social settings while ensuring a sufficient number of observations. The number of clusters per neighborhood varies widely, ranging from 1 to 234, depending on cluster density. We have opted for a 50km radius to have as many observations as possible of nearby women on the one hand, and on the other hand, making it likely that the involved clusters are in contact with each other. For robustness analyses, we added models in which we changed the range of the neighborhood into a smaller 20km radius and four more distant neighborhood areas: 75-90km, 100-112km, 125-135, and 150-158km[6].

To measure the variation in the timing of family formation for the three distinct social groups, we calculate the coefficient of variation (CV) for age at first marriage and age at first birth within each group. An advantage of CV is that the degree of variation is measured relative to the mean of the group. In other words, it takes into account that groups with higher means are likely to have higher standard deviations. The following standard formula is used to calculate the CV:

$$CV_{ij} = \frac{Std_{ij}}{Mean_{ij}} \qquad (3.3)$$

where *j* stands for the local neighborhood, ethnic, or religious group, respectively. We impose $N_j \geq 10$ to ensure that a meaningful CV can be calculated.

6. Except for the smaller 20km, for the four distant neighborhood areas we have chosen to keep the total surface of the neighborhood equal to the 50km area to make the results comparable.

Because social norms are likely shaped by older peers and prior cohorts, we calculate the CV solely on women who are of the same age or older than each respondent. The logic is that only these women could have plausibly shaped the norms and expectations relevant for the respondent's family formation decisions. To examine whether norms are shaped primarily by peers or by older generations, we further construct a series of age-specific variation measures. Specifically, we recalculate the CV for each respondent using reference groups that are 1–10 years older, 2–11 years older, and so on, up to 15–24 years older. Each of these specifications defines a separate estimation sample. We use 10-year age bands to have a sufficient number of observations in each group to calculate a CV on. Comparing the predictive strength of these different age-based measures allows us to assess whether social pressure operates more strongly through peers or through inter-generational transmission.

This age-stratified approach complements our broader sampling strategy which assesses to what extent social pressure influences women's timing of marriage and childbirth. For analyses of age at marriage and age at first birth, we restrict the sample to women aged 30 to 39 years who are currently or previously married and who have had at least one child. The lower age limit of 30 years is chosen so that we can reasonably argue that women who will get married and give birth have already done so in our sample[7]. The upper age limit of 39 years is chosen so that we can use the behavior of older women (women until 49 are interviewed in the survey) to calculate the CV in the timing of family formation. For analyses of the incidence of child marriage and teen pregnancy, we use a broader sample of all women aged 20 to 39, including those who have not married or given birth, since the incidence is fully determined by age 20.

All models control for the respondent's year of birth, years of education, and place of residence (urban or rural). We also include the number of women in each group used to compute the variation measures and the number of clusters that make up a neighborhood[8].

---

7. We see in our data that 95.0% of married women aged 40 to 49 married before the age of 30, and that 96.1% of the mothers in our data aged 40 to 49 gave birth to their first child before the age of 30.

8. A woman's wealth is not included in the models presented here, because a woman's contemporaneous wealth is also in part the result of the custom of dowry and is thus not representative of her situation at the age of marriage or giving birth to her first child. For completeness and to address concerns about omitted variable bias, in the appendix we present the models where wealth is included. They can be found in Tables A3.12:A3.15. These models show similar results as the ones presented in the main text in terms of effect sizes and significance.

Table 3.1 summarizes descriptive statistics for the two primary samples used in the analysis. Among ever-married women aged 30 to 39 years with at least one child, the mean age at first marriage is 18.8 years, and the mean age at first birth is 19.8 years. In the broader sample of women aged 20 to 39, 41% were married before age 18, and 51% gave birth before age 20 which reflects the persistent prevalence of early marriage and adolescent childbearing across the region[9]. We further observe that the average coefficient of variation (CV) in age at first marriage across neighborhoods, ethnic, and religious groups ranges from 0.223 to 0.235, while the average CV in age at first birth is slightly lower, between 0.200 and 0.209. This suggests that normative expectations around the timing of marriage are more variable across contexts than those related to childbearing.

Table 3.1: Descriptive statistics of the main datasets

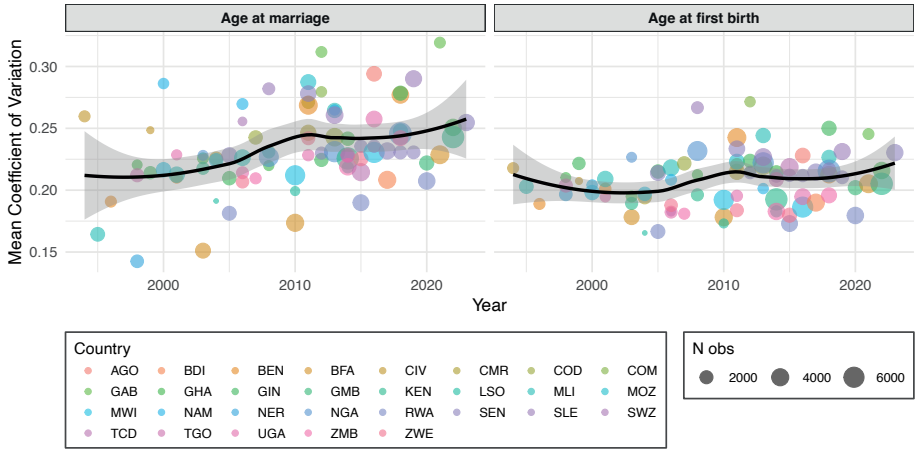
	Sample 1: Married women with children, aged 30-39		*Sample 2: All women, aged 20-39*	
	Mean / probability	**SD**	**Mean / probability**	**SD**
Age at first marriage	18.815	4.542		
Age at first childbirth	19.763	4.078		
Prob. child marriage			0.413	
Prob. teen pregnancy			0.505	
CV in neighborhoods age marriage	0.234	0.045	0.224	0.042
CV in religious groups age marriage	0.235	0.042	0.225	0.040
CV in ethnic group age marriage	0.232	0.044	0.223	0.042
CV in neighborhoods age first birth	0.208	0.031	0.201	0.030
CV in religious groups age first birth	0.209	0.030	0.201	0.029
CV in ethnic groups age first birth	0.207	0.031	0.200	0.030
Birth year	1978	7.249	1983	8.590
Place of living (urban)	0.343		0.368	
Years of education	4.006	4.594	4.802	4.851
Total observations	224,806		548,428	

9. Figure A3.1 in the appendix shows that the age at first marriage and first birth is increasing over time, and that the probability of being a child bride and teen mom is decreasing. Around 1995, the mean age of our sample for getting married is 17 and of giving birth is 18, whereas in 2022, this is almost 20 in both cases. If we look at child marriage, we see that probabilities decline from above 50% in 1995 to 30% in 2022. For teen pregnancy it declines from 60% in 1995 to almost 40% in 2022.

Figure 3.1 plots the evolution of the average CVs across all three social groups (neighborhoods, ethnic and religious groups) and for every DHS showing the developments over time. Confidence bands are indicated in grey and DHS-level averages are presented as circles with the size reflecting the underlying sample size. The CV in age at marriage has risen steadily since the late 1990s, from roughly 0.21 before 2000 to 0.25 after 2020, indicating a gradual loosening of social norms around the timing of marriage in recent decades. In contrast, the CV in age at first birth has remained comparatively stable over time, fluctuating narrowly between 0.20 and 0.22. This divergence hints at a differential pace of normative change: while norms around the timing of marriage appear to be relaxing, norms concerning fertility seem to remain more rigid.

Substantial cross-country variation underlies these trends. For instance, Burkina Faso (2003) and Niger (1998) display some of the lowest levels of variation in age at marriage (CV = 0.15), whereas Gabon (2012 and 2021) exhibits CVs above 0.30, which is twice as high. Similarly, in terms of variation in age at first birth, Lesotho (2004) and Rwanda (2005) show the lowest CVs (0.16), while Comoros (2012) and Sierra Leone (2008) have among the highest (0.27). These differences underline the importance of contextualizing social pressure within specific countries and cultural settings.

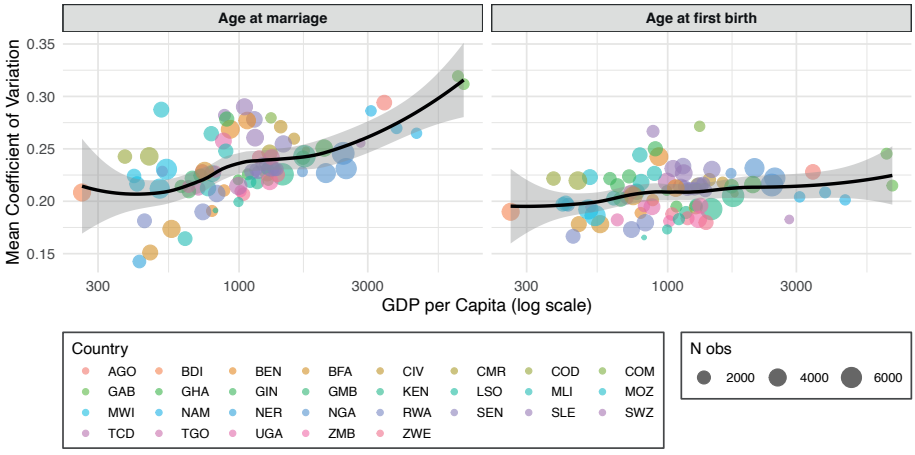
Figure 3.1: Development of mean CV over time



Notes: Each point represents the average coefficient of variation (CV) in age at first marriage and age at first birth across all neighborhoods, ethnic, and religious groups within a given DHS survey. CVs are based on women aged 30–39 for marriage and childbirth timing.

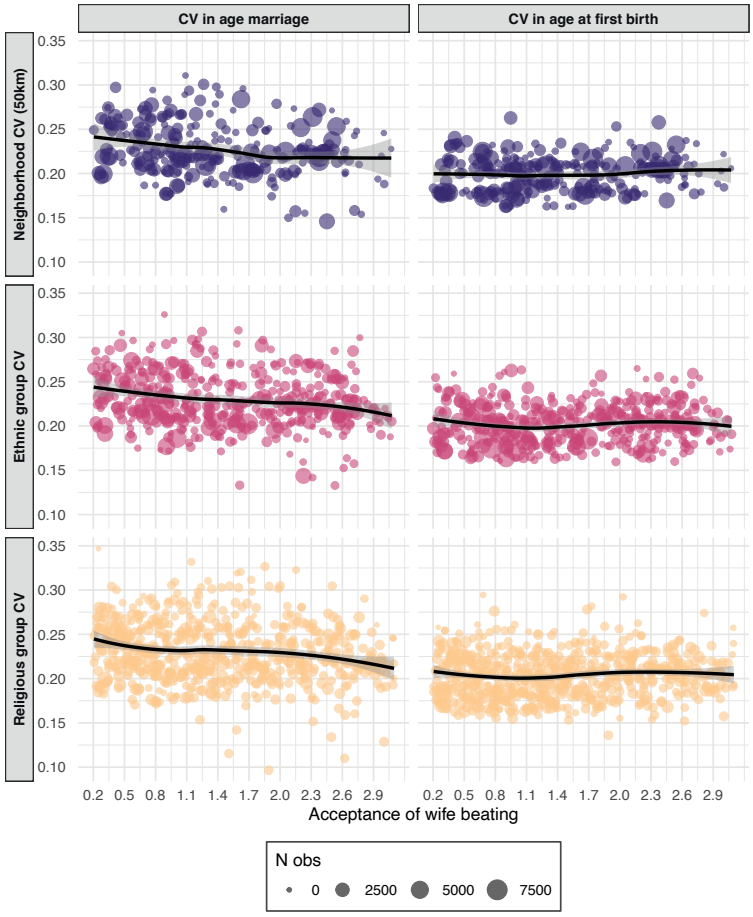
We also looked at the relationship between normative variation in the form of the average CV and a country's GDP to address to what extent economic development is associated with normative flexibility. Figure 3.2 shows that economic development alone cannot explain normative variation. While there is a positive correlation between a country's GDP and the average CV in age at marriage this relationship is not uniform as can be seen by the wide spread of circles representing CVs at every distinct level of income, i.e., the vertical spread. Countries with similar income levels often display strikingly different degrees of behavioral variation. For example, Mozambique (2011) and Burkina Faso (2003) had comparable GDPs at the moment of the respective survey of around $500 USD (in PPP), yet their CVs in age at marriage differ substantially. Normative variation ranks at 0.29 and 0.15 for Mozambique and Burkina, respectively. This underscores the importance of cultural and group-specific dynamics in shaping family formation trajectories. While there is also a moderate increase in CVs for age at first birth along the income dimension, this relationship is less pronounced, suggesting that normative variation around first birth is even less associated with a country's overall economic development. These figures further underscore the need to analyze the role of social norms as they have been fairly stable in the sample under study in the last 30 years (even the increase in normative variation in the age of marriage is only in the range of 0.04).

Figure 3.2: Relation between GDP and mean CV



Notes: Each point represents the mean coefficient of variation (CV) in age at first marriage and age at first birth across all neighborhoods, ethnic, and religious groups within a given DHS survey. CVs are based on women aged 30–39. GDP data reflect current US dollars and are drawn from the World Bank's World Development Indicators (https://data.worldbank.org/indicator/NY.GDP.MKTP.CD). The CVs and GDP are paired based on the year the DHS survey took place.

Figure 3.3: The relation between the coefficient of variation and acceptance of wife beating



Notes: Each point shows the relationship between the mean coefficient of variation (CV) in age at first marriage (left column) or age at first birth (right column) and the average acceptance of wife beating, across neighborhoods (top row), ethnic groups (middle row), and religious groups (bottom row). Acceptance of wife beating is measured as the mean score across five DHS items asking whether it is justified when a wife: (i) argues with her husband, (ii) goes out without telling him, (iii) burns the food, (iv) refuses sex, or (v) neglects the children (0 = no, 1 = yes). CVs and attitudes are calculated using the same sample of women aged 30–39.

We further explore how the CV relates to broader gender-related normative attitudes by examining its correlation with the acceptance of wife beating. Figure 3.3 shows a very modest negative relationship between the CV in age at marriage within neighborhoods, ethnic groups, and religious groups, and the average acceptance of wife beating. This suggests that in groups where there is greater behavioral diversity

in marriage timing – and thus weaker normative enforcement – respondents are less likely to find wife beating acceptable. By contrast, we find no clear relationship between variation in age at first childbirth and attitudes toward wife beating. This divergence may suggest that variation in the timing of marriage is more closely aligned with attitudes toward gender norms, whereas norms around the timing of fertility are shaped by different factors.

Thus, the descriptive analysis of normative variation in SSA allows us to draw four initial conclusions: (i) there is huge variation in norm strength across contexts, (ii) in the last 30 years the normative commitment related to the timing of marriage decreased while it stayed stable for fertility timing, (iii) normative variations are at most moderately associated with a country's overall economic development, and finally (iv) normative variation in the timing of marriage tends to show different associations with external factors compared to normative variation in the timing of the first child stressing that norms around marriage differ from those related to fertility.

While these are global associations, next, we want to zoom in on the micro-level and assess to what extent normative variation at the group level affects individual level decision making.

3.5 Empirical framework

To examine the relationship between cultural variation and an individual woman's timing in family formation, we use a fixed effects regression framework. The model is specified as follows:

$$Y_{ij} = \beta_0 + \beta_k CV_{ijk} + \beta_x X'_i + \rho_j + \tau_s + \zeta_{by} + \sigma_c + \lambda_r + \eta_{icjt} \quad (3.4)$$

where Y_{ij} is the outcome variable, which is in our case one of the following four variables: (a) the age at marriage, (b) the incidence of child marriage, (c) the age at giving first birth, and (d) the incidence of teen pregnancy. We have information for woman *i* at subnational region *j*. Our relationship of interest is represented by CV_{ijk} with $k = n, e, r$. It captures the variation in family formation within a woman's neighborhood if $k = n$, a woman's ethnic group if $k = e$, and a woman's religious group if $k = r$. In addition, we control for individual women specific co-variates, X'_i, which are a woman's place of living and years of education. We also include

fixed effects for subnational regions (province) denoted by ρ_j, survey denoted by τ_s, birth year of the woman denoted by ζ_{by}, and the ethnic and religious group of the woman, denoted respectively by σ_e and λ_r. Standard errors are clustered at the region level.

We also tested whether the effect of CV_{ijk} with $k = n, e, r$ differs between women with a different background regarding years of education and place of living by adding interaction terms between CV_{ijk} and years of education as well as place of living. To address the robustness of our findings, we added models where we tested different ranges of the neighborhood, we controlled for influential countries, and we tested whether household wealth is driving our results.

Since we rely on the pre-determined, realized community variation in marriage and family formation in our main approach, we consider the estimated β coefficients associated with the different coefficient's of variation as conditionally causal, conditional on the control variables and the fixed effects that we include.

3.6 Results

3.6.1 The influence of 'normative variation' across communities on the timing of family formation

We organized the presentation of our main results graphically in Figure 3.4. The figure contains four graphs. Each graph presents the results for one of the four outcomes: (a) age at marriage, (b) incidence of child marriage, (c) age at first birth, and (d) incidence of teen pregnancy. Within every graph, we present the results associated with the effect of diversity in a woman's different social groups that stem from four different models. The coefficient estimates denoted by 'n' (1n, 2n etc.) represent the impact of diversity in local neighborhoods, those denoted by 'e' refer to diversity in ethnic groups, and those indicated with 'r' consider diversity in religious groups. Coefficient estimates denoted by 'all' are derived from specifications that jointly include all three diversity measures. Full regression results for all models can be found in Appendices B A3.2 to A3.5.

Even at a glance it can be seen that the results show a postponement of family formation due to norm variation in a woman's social groups. Model 1n presented in Panel A of Figure 3.4 shows that women tend to marry later when their local neighborhoods have greater variation in age at marriage as measured by the CV. Women living in neighborhoods with a variation of one standard deviation (SD)

above the mean marry 1.6 months (= 0.13 years) later than those in neighborhoods with average variation. Models 1e and 1r show similar results for ethnic and religious groups. Women living in ethnic and religious groups with a variation of one SD above the mean marry 1.4 months (= 0.12 years) later on average than those in groups with average variation. These findings support the theoretically posited expectation that women in more diverse social groups tend to exhibit less traditional family formation behavior. When the three variation measures (neighborhoods, ethnic, and religious groups) are combined in one model (1all), the regression coefficients decrease in magnitude but remain significant for all three social groups. Women living in neighborhoods, ethnic, and religious groups with a variation of one SD above the mean marry on average 2.6 months (= 0.22 years) later compared to women in groups with average variation. Given the average marriage age of 18.8 years in the sample, this corresponds to a 1.2% increase in age at marriage per standard deviation of variation. If we compare this effect of group variation with the effect of education (Appendix B Table A3.2), we see that with 1 additional year of education, women marry 3.6 months (=0.30 years) later. As such, the effect of 1 SD variation in CV is slightly less compared to the effect of an additional year of education but still sizable as it corresponds to 72% of the education effect but in most if not all of the existing quantitative marriage models it is captured as unobserved.

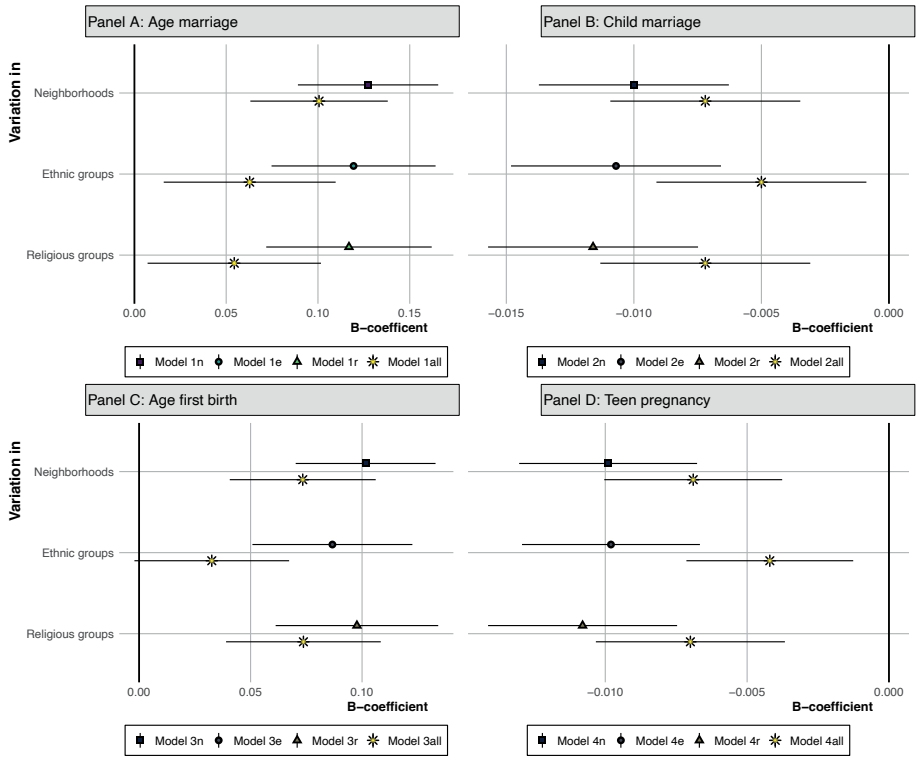
Models 2n, 2e and 2r in Panel B replicate the patterns observed in the models indexed with 1, but focus on a woman's incidence of child marriage. Women living in neighborhoods, ethnic, and religious groups with a variation of one SD above the mean have a 1.0, 1.1 and 1.2 percentage points lower chance of becoming a child bride, respectively. Considering the three social groups together in model 2all, we find that women living in groups with a 1SD above the mean variation have a combined lower probability of child marriage of 1.9 percentage points. Given the sample mean of 41% for child marriage, this implies a 4.7% reduction in the chance of child marriage for women living in groups with higher variation (+1 SD).

Panel C shows the results for a woman's age at first birth and her respective social groups' variation in age at first birth. Women in neighborhoods with higher variation give birth 1.2 months (= 0.10 years) later on average, while those in more diverse ethnic and religious groups give birth 1.0 months (= 0.09 years) and 1.2 months (= 0.10 years) later, respectively. When all three social groups are considered together, the delay in age at first birth increases to 2.1 months (= 0.18 years). The variation in neighborhoods has the most significant impact on age at first birth, while ethnic group variation becomes insignificant. If we compare women from social groups with higher variation

(+1 SD) to women from average variation, we observe a 0.9% increase in age at first birth given the sample mean of 19.8 years.

Models 4 in Panel D assess the likelihood of teen pregnancy. Women in more diverse neighborhoods, ethnic, and religious groups have respectively a 1.0, 1.0, and 1.1 percentage points lower chance of being a teen mom. When all three groups are considered, the chance of teen pregnancy decreases by 1.8 percentage points. This corresponds to a 3.6% reduction when comparing women living in groups with 1 SD above versus the average variation, given the sample mean of 50.5% for teen pregnancy.

Figure 3.4: Effect of cultural diversity on women's timing of family formation



Notes: The figure presents regression estimates and 95% confidence intervals on the x-axis for the effect of normative variation as measured by the coefficient of variation (CV) on four outcomes: (a) age at first marriage, (b) incidence of child marriage (before age 18), (c) age at first birth, and (d) incidence of teen pregnancy (before age 20). Results are shown separately for variation within neighborhoods ('n'), ethnic groups ('e'), and religious groups ('r'), as well as the joint effect of including all three CV measures together in the same model ('all'). All models include country-year fixed effects, birth-year fixed effects, ethnic group fixed effects, and religious group fixed effects together with the individuals' years of education, place of living as well as the number of clusters and individuals based on whom the variation measurement is calculated. The full regression results can be found in Appendices B A3.2 to A3.5.

Overall, these findings show that increased variation in a woman's surrounding social environments – neighborhoods, ethnic, and religious groups – is associated with less traditional family formation trajectories. The influence is manifested in delayed marriage and childbirth (1-3 months) and a reduced likelihood of becoming a child bride or teen mom (1-2 percentage points) per standard deviation in the CV. Moreover, we show that these results are robust to different neighborhood sizes, but that the effect is strongest for the 0-50km neighborhood (see Figure A5)[10]. The results are also robust to the exclusion of individual countries (see Figures A6 to A9), and to the inclusion of contemporaneous household wealth (see Tables A12 to A15).

3.6.2 The influence of 'normative variation' across cohorts

Figure 3.5 presents the regression estimates showing the effect of 'normative variation' across cohorts on a woman's timing of family formation. We move from the cohort of peers that is 1-10 years older to the cohort of 2-11 years older individuals to the 3-12 years older cohort. We let the cohorts age by one year in going from one to the next. The oldest cohort considered is 8-17 years older. We implement an age range of 10 years per cohort to ensure that we have enough observations to credibly calculate the CVs. As with the previous figures we present four panels according to our four outcome variables. Yet, since we introduce the time dimension here and we get one estimate per cohort, we split up the presentation of the effects into three subgraphs by neighborhood, ethnicity and religion. Confidence intervals are shown along with the point estimate for every cohort estimate. The exact β-coefficients can be found in Appendices B A3.6 to A3.11.

Panel A of figure 5 shows the effects of normative variation in the different age cohorts on a woman's age at marriage. As can be seen from the decreasing relationship presented in the first graph (model 5n), the normative variation in the neighborhood community has the most impact when the reference group is within a similar age range. As the age gap between the woman and the reference group increases, the influence diminishes in strength. Specifically, the regression coefficients for the normative variation in a woman's neighborhood (model 5n) decreases by half, from a significant 0.076 for the age cohort that is 1-10 years older to an insignificant 0.037 for the cohort that is 8-17 years older. This trend is less pronounced for the normative variation in ethnic and religious groups (respectively models 5e and 5r), where the effect of normative variation on the timing of marriage remains significant across age cohorts. This suggests that social pressure from

10. The most distant range of the neighborhood that we tried is 150-158km. Moving further away from the place of residence of the women is difficult regarding the country borders of the smaller countries in our sample.

older generations may operate primarily through a woman's ethnic and religious background rather than through her immediate neighborhood.

Models 6n-6r of panel B present similar findings for the incidence of child marriage. Unlike the previous models, we find that the normative variation in social groups remains significant regardless of the age cohort sub-sample for all three social groups. Although the regression coefficients halve for the neighborhood and more than halves for ethnic and religious groups when going from the oldest cohort to the youngest, normative variation within these social groups continues to significantly predict a woman's likelihood of becoming a child bride. This suggest that the norms to comply with the practice of child marriage are still present within the younger generation with direct impacts on the lived reality of young women, giving some additional insights why the practice of child marriage is difficult to change. At the same time, the results show how contemporaneous decisions about early marriage are largely influenced by the norms of older generations, potentially independent of a concerned woman's own preferences.

Models 7n-7r of panel C introduce the findings associated with the effects of normative variation on a woman's age at first childbirth. As in the previous models, the regression coefficients are strongest when the social groups are more similar in age to the woman. However, a key difference is that the coefficients for ethnic and religious groups become, as for the neighborhood, insignificant as the age gap between the woman and the reference group increases. This suggests that social pressure to conform to norms around early childbirth may primarily come from peers, rather than from older generations.

Lastly, models 8n-8r of panel D examine the effect of normative variation in age at first childbirth on a woman's likelihood of teen pregnancy. In line with the results from Models 7n-r on age at first birth, we find that the influence of all three social groups —neighborhoods, ethnic groups, and religious groups — becomes insignificant as the age gap between the woman and the reference group increases. For neighborhoods, the regression coefficient becomes insignificant when the reference group is 11 to 20 years older, whereas for ethnic and religious groups, this shift occurs slightly later, with the age cohort of 14 to 23 and 13 to 22 years, respectively.

Overall, the results from models 5-8 show that variation within social groups has a stronger influence on a woman's timing of family formation when the social group is closer in age to her. Regarding the timing of marriage (both age at marriage and

child marriage), we find that previous generations play a more significant role in exerting social pressure to conform to the norm of early marriage. Particularly for the extreme form of early marriage, child marriage, this social pressure persists across generations. This is in line with existing evidence that (child) marriage is often arranged by parents and other older family and community members, keeping the influence of older generations high. In contrast, for the timing of childbirth (age at first birth and teen pregnancy), we find a different pattern, as the normative influence of previous generations on contemporaneous women's decisions becomes insignificant. As such, social pressure to conform to norms around early childbirth seems to be more likely to come from peers rather than parents and older community members.

3.6.3 Heterogeneity analyses: The role of education and place of living

We also tested whether the delaying effects of normative variation in social groups differs by a woman's individual characteristics, like her level of education and place of living.

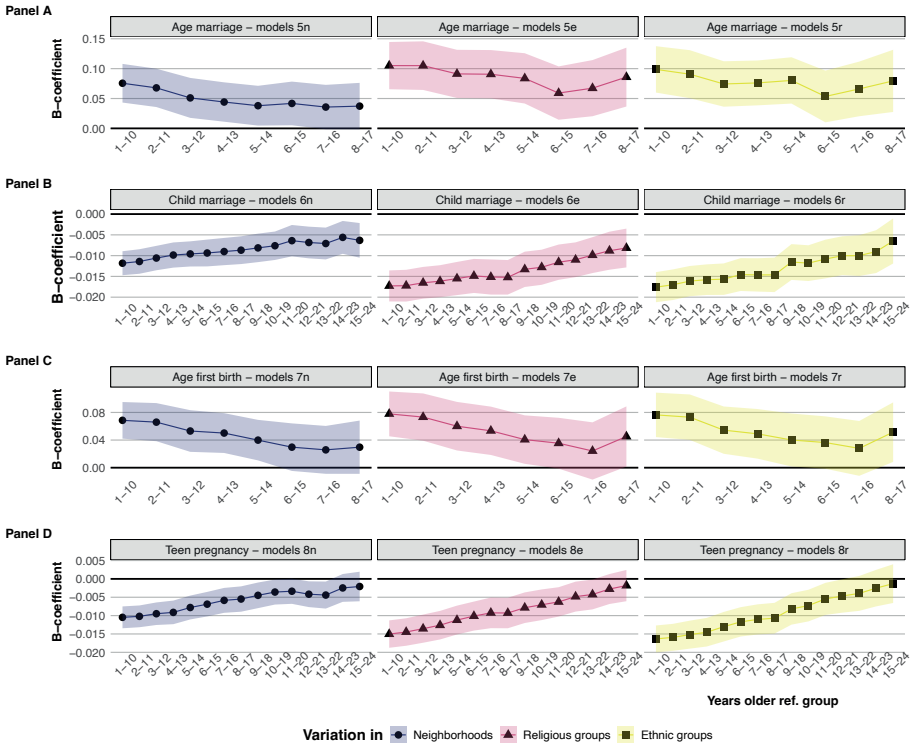
For a woman's age at marriage, we find no significant interaction between behavioral variation in social groups and either education or urban residence (see figure A2). This suggests that normative pressure, as proxied by group-level behavioral variation, affects women similarly across these characteristics when it comes to general timing of marriage.

However, this changes when we examine the incidence of child marriage, a more extreme form of early marriage behavior. Figure 3.6 shows predicted probabilities of child marriage across different levels of normative variation within the three reference groups: neighborhoods, ethnic and religious groups. The top row (Panels A-C) shows the predicted probabilities for women with 3 different levels of education [0, 4, and 8 years]. The bottom row (Panels D-F) shows the same for a woman's place of residence, distinguishing between rural and urban residence.

Panels A-C show that there is a significant interaction between education and normative variation. Among women with no formal education, higher variation in group-level age at marriage significantly reduces the likelihood of becoming a child bride. This corresponds to our earlier findings and supports the interpretation that greater behavioral diversity weakens the pressure to conform. For women with four years of education (the sample median), the effect remains negative but is weaker. If a woman has 8 years of education, there is no longer an effect of the CV on a woman's probability of child marriage. This pattern suggests that education

may serve as a protective factor against social pressure, reducing the influence of community norms on early marriage decisions.

Figure 3.5: Effect of variation on women's timing in family formation

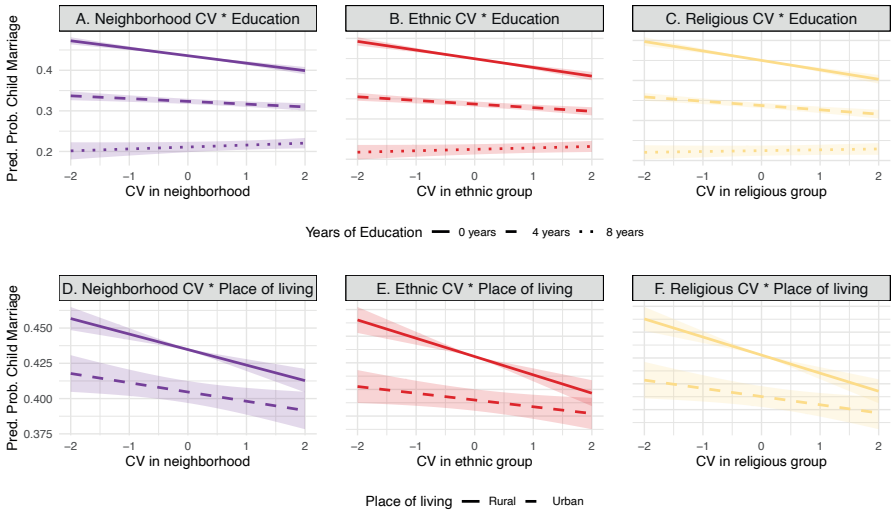


Notes: The figure presents regression estimates and 95% confidence intervals for the effect of the normative variation measured by the standardized coefficient of variation (CV) on four outcomes: (a) age at first marriage, (b) incidence of child marriage (before age 18), (c) age at first birth, and (d) incidence of teen pregnancy (before age 20). Each point represents the effect of CVs calculated over different age bands within a woman's social groups starting from the cohort of those being 1 to 10 years older to the cohort of those being up to 15 to 24 years older. Results are shown separately for variation within neighborhoods ('n'), ethnic groups ('e'), and religious groups ('r'). All models include country-year fixed effects, birth-year fixed effects, ethnic group fixed effects, and religious group fixed effects together with the individuals' years of education, place of living as well as the number of clusters and individuals based on whom the variation measurement is calculated. The exact β-coefficients can be found in Appendices A3.6 to A3.11.

Panels D–F show a similar interaction with the place of living. The negative effect of normative variation on child marriage is significantly stronger in rural areas compared to urban areas. However, even in urban settings, the effect of the CV remains negative, indicating that social pressure persists, though it is attenuated in more diverse or cosmopolitan contexts.

We observe comparable patterns for age at first birth and teen pregnancy (Figures A3 and A4). A woman's level of education and place of living do not moderate the effect of the CV on age at first birth, but they do mitigate its effect on teen pregnancy. Specifically, the negative relationship between the CV and teen pregnancy disappears among women with at least eight years of education. For urban women, the effect is also weaker compared to their rural counterparts.

Figure 3.6: Interaction effects between CV, education and place of living on child marriage



Notes: The figure shows the predicted probabilities of child marriage across different levels of normative variation (measured by the standardized coefficient of variation, CV, in age at marriage within the reference group) and the 95% confidence interval. The top row (Panels A–C) plots the predicted probabilities separately for neighborhoods (A), ethnic groups (B), and religious groups (C) for 3 levels of education: 0, 4 and 8 years. The bottom row (Panels D–F) shows the same for the place of residence, distinguishing between rural and urban women. All models include region fixed effects, birth year fixed effects, survey fixed effects, and ethnic and religious group fixed effects, as well as controls for years of education, urban residence, and the number of individuals and clusters on which the variation is computed.

These patterns indicate that while some norms around early family formation remain deeply rooted, their enforcement is not fixed. Education and urban living, which both expand a woman's exposure to alternative behaviors and networks, appear to weaken the conformity pressures associated with child marriage and teen pregnancy. In contrast, the general patterns of early marriage and childbearing are less sensitive to these factors, perhaps because they are more socially acceptable and less strictly enforced by community norms.

3.7 Discussion and conclusion

The aim of the present study is to examine whether variation in family formation trajectories within social groups can serve as a proxy for social pressure to conform to cultural norms in SSA. Drawing on the concept of the social image utility (Bursztyn & Jensen, 2017), we argue that in contexts where behavior is highly uniform, social norms are likely strong and deviations are more heavily sanctioned. In contrast, greater behavioral diversity implies weaker enforcement of norms and increased scope for individual decision-making. We empirically operationalize social norm strength with a standardized measure of variation – the coefficient of variation (CV). We find robust evidence that higher within-group variation in the timing of marriage and childbirth is associated with later transitions to marriage and motherhood and a lower likelihood of becoming a child bride or teen mother.

These findings suggest that diversity in the timing of family formation reflects a relaxation of social expectations and a loosening of traditional norms. In more diverse groups, women appear to face less pressure to conform to culturally sanctioned timelines, allowing their decisions to be shaped more by individual preferences and individually perceived costs and benefits (Field & Ambrus, 2008; Hicks & Hicks, 2019). Importantly, our analysis shows that these patterns hold across different types of social groups –neighborhoods, ethnic, and religious groups – highlighting that variation is not merely geographically rooted but deeply embedded in social identity structures. Moreover, the persistence of effects across generations for marriage outcomes, but not for childbirth, suggests that norms around marriage are more strongly transmitted from older to younger cohorts, whereas norms around fertility may be more contemporaneously shaped by peers. In addition, the effects of normative variation on child marriage and teen pregnancy are significantly stronger among women with little or no education and those living in rural areas. Among more educated and urban women, these effects weaken substantially, suggesting that individual resources and broader social exposure mitigate the influence of group-level social pressure.

By leveraging cross-sectional data from 104 DHS surveys across 33 SSA countries, our approach offers a scalable and policy-relevant way to measure the localized strength of cultural norms. Whereas previous studies often rely on national-level indicators, such as the cultural tightness-looseness index (Fischer, 2009; Gelfand et al., 2011), or qualitative assessments that are rich in context but difficult to generalize (Ahonsi et al., 2019; Mkhwanazi, 2010), our method captures norm enforcement at the level of any social group that shapes individual behavior. This allows for a more nuanced understanding of how cultural traditions persist and why behavior change has been slow in certain contexts, despite increasing access to education, wealth, and legal reforms (Batyra et al., 2021; Castro Torres et al., 2022; Jensen & Thornton, 2003). Most importantly, our measure of normative variation does not suffer from social desirability bias as we are not asking any questions related to normative judgments but make use of location, ethnic, and religious affiliation. This is a key contribution to the existing literature on social norms. Our results also help explain why policy interventions have had mixed effects in SSA (Feyissa et al., 2023): communities with the strongest normative expectations are often those most resistant to change, yet they are precisely the ones most in need of support.

Nonetheless, the study has some limitations. First, we are unable to observe whether the women under study still reside in the communities where they spent their formative years. Migration may attenuate the relationship between current social context and earlier life events. This issue is most relevant for neighborhood-level indicators but less so for ethnic and religious affiliations, which are generally stable over time. Another limitation is that we assume early marriage and childbirth to represent traditional behavior across all contexts, while it is possible that some groups may have strong norms favoring later transitions. To understand such group level differences in perceptions, it would require a more nuanced, group-specific conceptualization of tradition. This is beyond the scope of the study at hand but an important avenue for future research. Third, our measure of variation is based on a minimum of ten observations per group, which may introduce some imprecision, particularly in sparsely populated areas. Finally, our empirical framework does not capture the extent to which individuals internalize or identify with group norms, which is a key component of social image concerns. Future research could explore the role of norm salience, reference group salience, or perceived sanctions for non-conformity.

Taken together, our findings highlight the need to look beyond formal policies and individual attributes to understand how social norms continue to shape early

family formation in SSA. By introducing a measurable and context-specific proxy for social pressure, this study offers new insights into the mechanisms through which cultural norms influence behavior –even in the face of growing access to education, economic opportunities, and legal protections. As efforts to reduce child marriage and teen pregnancy continue, it becomes increasingly important to recognize and address the role of group-level expectations and social influences. Identifying where pressure to conform remains strongest is not only key to explaining the limited success of past interventions, but also essential for designing more effective, locally adequate, and culturally responsive strategies in the future.

Chapter 4: Fertility rates and social media coverage in sub-Saharan Africa[11]

[11.] This chapter is based on joint work with Jeroen Smits and Sandor Schrijner and has been published as:

Wildeman, J., Schrijner, S., & Smits, J. (2023). Fertility rates and social media usage in Sub-Saharan Africa. *Population, Space and Place, 29, e2635*. https://doi.org/10.1002/psp.2635 The study on which this chapter is based was presented at the European Association of Population Studies Conference 2022.

This study examines the connection between social media coverage and fertility behavior in sub-Saharan Africa (SSA). Where the link between traditional media and fertility behavior has been firmly established, the link between social media, which has expanded enormously in recent years and provides opportunities for information and social interactions, and fertility is yet unclear. We combine digital trace data from Facebook with fertility rates for 311 subnational regions in 29 sub-Sahara African countries. Our analysis starts with describing spatial patterns of crude birth rates, Facebook coverage and gender gaps in Facebook coverage and continues with regression analysis to test whether the associations remain significant after controlling for confounders. We find higher social media coverage and gender equality in social media coverage to be associated with lower birth rates. This study is one of the first to demonstrate connections between social media coverage and fertility outcomes in the SSA region.

4.1 Introduction

Fertility levels in sub-Saharan Africa (SSA) are known to be influenced by traditional media exposure, like watching television and listening to the radio (Ferrara et al., 2012; Inkeles & Smith, 1974; Lerner, 1958; Westoff et al., 1997). Since 2000, rapid expansion of information and communication technology (ICT) -- often referred to as the digital revolution -- has caused a major shift in the availability and role of media. In 2020, 93% of the Africans had access to a mobile phone, whereas this was only 53% in 2011, and 30% of the Africans were connected to the Internet in 2020 (World Bank, 2021b). A key characteristic of media after the digital revolution is digital interactivity between media consumers, which is predominantly present in social media. This raises the important question to what extent social media coverage is associated with fertility levels.

A main determinant of SSA's fertility levels is women's education, acting through several mechanisms such as supporting women's empowerment, facilitating their labour force participation, and reducing child mortality. Another determinant of fertility levels are interactions between individuals within communities, which may lead to the diffusion of information, ideas, and technologies regarding for example contraceptive usage, fertility preferences and family size (Bongaarts & Watkins, 1996; Coale & Watkins, 1986; Lesthaeghe, 1978). Individuals receive and interpret this diffused information differently. The information may therefore gain new meaning in different social contexts and ultimately promote or constrain fertility behaviour and outcomes. Important places were diffusion takes place is in schools (Atake & Gnakou Ali, 2019; Bongaarts, 2020; Shapiro & Tenikue, 2017). Diffusion takes also place by traditional media (Ferrara et al., 2012; Westoff & Koffman, 2011).

Compared to traditional media, social media offer additional opportunities for the diffusion of ideas. Social media transcend country borders more easily and thus promotes exposure to less context-specific and more globalized cultural ideas, including ideas about fertility behavior. Access to social media provides access to the "life of others" without being reliant on physical proximity. Moreover, social media enable their users not only to consume, but also to contribute content and become member of an online community of contributors (Kietzmann et al., 2011). As such, social media provide individuals the possibility to obtain new information and to interact with role models beyond their own community.

Earlier studies suggest that the impact of the digital media, like social media, is likely to be larger in developing regions with poor other types of infrastructure.

Due to social media, individuals living in remote areas can get in contact with new ideas and individuals beyond their community borders, from which they were disconnected before due to a lack of infrastructure (Aker & Mbiti, 2010; Rotondi et al., 2020). Studies also showed that within these remote areas, the impact of the digital revolution is found to be larger for more marginalized groups, like females who face greater barriers to access digital media (Rotondi et al., 2020; Varriale et al., 2022).

Due to the difficulty to measure social media access, previous research predominantly relied on national estimates of official ICT statistics. These estimates potentially mask underlying heterogeneity within countries. The strength of the current study is that it uses new digital trace data of social media, in particular Facebook, at the subnational level to uncover these underlying differences. Digital trace data has its own advantages, like the high speed of large-scale data collection and the low costs that are involved (Cesare et al., 2018). To resolve the potential lack of accessibility, transparency and representativeness of this type of data, we will enrich it with traditional survey data. This approach has already yielded promising results in previous studies with regard to data reliability (Eker et al., 2021; Kashyap et al., 2020; Rampazzo et al., 2021; Ribeiro et al., 2020).

This study will use subnational digital trace data from Facebook -- one of the largest social media platforms in SSA -- to examine to what extent social media coverage and gender gaps in coverage explain subnational regional variations in fertility levels. For our analyses, we have combined Facebook data with household survey data for 311 subnational regions in 29 SSA countries. This enables us to study the relationship between fertility and social media coverage in more detail than has been possible before.

In the next section, we provide an overview of the relevant literature concerning the relationships between fertility behaviour and (social) media. The result section describes the spatial pattern in subnational variation in birth rates, social media coverage and gender gaps in social media. Subsequently, the association between birth rates and social media coverage will be tested using regression models.

4.2 Background

4.2.1 A history of media exposure and fertility behaviour

There is a long history of studies that examined the relationship between media exposure and fertility behavior. One of the first studies into this topic is from Lerner (1958), who reported that people who were more exposed to media (radio and film) were more likely to express "modern" and "pro-development" attitudes of various sorts in six countries in the Middle east. Some decades later, similar effects were found in studies focusing on SSA. Inkeles and Smith (1974) e.g. found that watching television, listening to the radio and reading newspapers were good predictors for "modern" norms and values, and this included the will to limit family sizes. Westoff and Bankole (1997) found a strong positive association between media use and fertility behavior in six SSA countries (Burkina Faso, Ghana, Kenya, Madagascar, Namibia, and Zambia). Included fertility behaviors were knowledge of modern contraception methods, intention to use contraceptives, current usage of contraceptives, and the desire to limit one's family size. An additional analysis of the same authors showed that women who were 'heavy media consumers' were more likely to use contraceptives and to limit their family size than women who did not consume media. However, caution should be in place when interpreting the causality of this relationship. According to the authors there is a possibility that heavy media consumers were already on a different reproductive trajectory before the study started, which could be the cause of their media consumption.

Another study in the developing world reported a causal link between media and fertility. La Ferrara, Chong and Duryea (2012) examined the effect of exposure to soap operas on fertility rates in Brazil. The main characters in most of the soap operas were living in urban settings and had lifestyles that were in stark contrast with those of people living in rural settings. In comparison with their public, the main characters in the soap operas had small families, married later than average, and had a high level of education. The study reported that the availability of a television signal, and thus exposure to soap operas, decreased the probability of giving birth to a child. This effect was the strongest for women who had a low socio-economic status and which lives were thus most in contrast with the lives of the main characters.

Media exposure, however, does not always show a positive relationship with globalized norms and values. For example, media exposure at least once a week was associated to lower acceptability of intimate partner violence in only 14 out of 26 low- and middle-income countries (Pierotti, 2013). Another study found a

positive association between media access and the odds of tolerating partner violence in Nigeria for women, but a negative association for males (Okenwa-Emegwa et al., 2016). A similar finding was reported by Banerjee, La Ferrara, and Orozco (2019). They found that exposure to MTV series was only associated with lower acceptance of partner violence among males and not among females in Nigeria. As such, media exposure can also be negatively associated with liberal and globalized fertility behaviors.

4.2.2 From traditional to social media

Since 2000, media evolved from 'old' or 'traditional' media to 'new' media, which is often referred to as the 'media revolution' or 'digital revolution'. Where traditional media predominantly consisted of listening to the radio, watching television and reading newspapers, new media are characterized by digitalization of the content. New media continuously evolve and consequently, the content of it changes frequently. In the beginning, new media were often considered as complementary to traditional media with individuals simply consuming additional online content, such as individuals passively reading and watching the online media (DiMaggio, Hargittai, Neuman, et al., 2001; Kietzmann et al., 2011).

One study that focused on the relationship between new media and fertility research showed that women's internet exposure was positively related to the usage of modern contraception in eight sub-Saharan countries. The increase in usage of modern contraception was largest among poorly educated women (Toffolutti et al., 2020). Another study examined how access to mobile phones influenced fertility behavior in Balaka, Malawi (Billari et al., 2020). This study found that the possession of a mobile phone was associated with smaller family size preferences. Moreover, the authors concluded that phone ownership may accelerate the fertility transition in SSA.

Social media have unique characteristics in comparison to digital media in general, such as the interactivity among its users. Using a narrow approach, social media involve person-to-person relationships on social networking services such as Twitter and Facebook. However, more often a broader interpretation of social media is used in which social media are described as interactive platforms through which individuals and groups can share, co-create, discuss, and modify self-created content (Kietzmann et al., 2011). As social interaction is known to be an important mechanism through which media influences fertility behaviour (Bongaarts & Watkins, 1996; Coale & Watkins, 1986; Ferrara et al., 2012; Lesthaeghe, 1978; Westoff & Koffman, 2011), the interactivity of social media is likely to influence fertility ideas and preferences. Online communities enable individuals to discuss ideas

and information with others without being reliant on physical proximity, which could lead to getting in contact with persons living in areas where other fertility patterns prevail.

One qualitative study examined the role of WhatsApp, which is a form of social media, among females living in a community in Nigeria (Abubakar & Dasuki, 2018). This study found that females experienced more opportunities to voice their personal opinion and autonomy due to getting in contact with other females through group chats in WhatsApp. Females were able to push the barriers of social norms that restricted them in gendered roles within their community. As such, this study showed that individuals who have access to online communities may be less dependent on societal norms and values of their physical surrounding and are more likely to get in contact with globalized norms and values. Hence, even though online communities are often formed by self-selection and homophily, they also provide users with the possibility to transcend the borders of their physical surrounding. As globalized norms are in favor of smaller family sizes than the average family sizes in SSA (Bongaarts, 2017), our first hypothesis is that social media coverage links to lower birth rates.

4.2.3 Inequalities in social media coverage

Another aspect of new and social media is that they are not universally within reach of everybody. There are significant inequalities in access to ICTs, which are often referred to as the 'digital divide' (DiMaggio, Hargittai, Neuman, et al., 2001; Norris, 2001). Known dimensions of the digital divide are level of development and gender. Access to new media in developing countries have been found to reflect broader structural inequalities, like inequalities regarding access to education, employment and income (Kashyap et al., 2020; L. Robinson et al., 2015). Moreover, studies reported that women are underrepresented in online populations of Google and Facebook in South Asia and in SSA (Fatehkia et al., 2018; Kashyap et al., 2020).

The digital gender divide is also associated with cultural factors (Scheerder et al., 2017). Cultural norms in patriarchal contexts may hinder females' access to new media, especially when this access is mediated through males (Abu-Shanab & Al-Jamal, 2015; Gurumurthy & Chami, 2014). Other studies show that females' adoption of new and social media is related to women's empowerment and societal well-being, and to globalized and liberal values (Abubakar & Dasuki, 2018; Varriale et al., 2022). Hence, our second hypothesis states that when gender gaps in access to social media are smaller fertility levels are expected to be lower.

4.3 Data

This study examines to what extent social media coverage and gender differences in social media coverage are related to fertility levels in subnational regions of SSA countries. We combine data from Facebook's marketing application with development indicators computed on the basis of household surveys. The household surveys are derived from three sources: Demographic and Health Surveys (DHS, www.dhsprogram.com), UNICEF Multiple Indicator Cluster Surveys (MICS, mics.unicef.org) and Afrobarometer Surveys (www.afrobarometer.org). These three sources are publicly and freely accessible for academic purposes.

Subnational regions are the unit of analysis in our study. The selection of subnational regions is based on the geographic coding available in the household survey datasets and the regions used by Facebook's marketing application. The available regions in the household surveys were taken as starting point to select regions used by Facebook's marketing application. We were able to make connections for 311 regions in 29 countries. An overview of the included countries, subnational regions and data sources is presented in Appendix 4.I.

4.3.1 Key fertility measure

The dependent variable is the crude birth rate (CBR) of the subnational regions in the year 2020. The CBR represents the number of yearly childbirths per 1000 individuals in a region. It has the great advantage over other fertility measures, like the total fertility rate (TFR), that it is a very simple measure that can be reliably computed over a short time period. As the spread of social media in the SSA region goes very rapidly (Nanfuka, 2022), it is essential that the measurement of our dependent variable is close in time to our measurement of social media coverage. Because the CBR might be influenced by the age structure of the female population, we have included the percentage of fertile women in a region's total population, and the mean age of the fertile women in a region as extra control factors in our models.

The subnational CBRs for 2020 were estimated by applying the subnational variation in CBRs derived from the most recent DHS and MICS surveys to the national CBRs for 2020 derived from the World Development Indicators (WDI) (World Bank, 2021a). To do this, we follow a procedure derived from (Kummu et al., 2018; Smits & Permanyer, 2019) in which subnational variation derived from household surveys is normalized around national values obtained. To do this, first the national crude birth rate from the World Bank in 2020 was divided by the national crude birth rate of the most recent DHS or MICS survey to create a differentiation factor. Second,

the subnational crude birth rates from the DHS/MICS surveys were multiplied with this factor to estimate the subnational crude birth rates for the year 2020. By doing this, we assume that the subnational crude birth rates follow the same pattern of change as the national values between the survey year and 2020.

Because the size of the estimation error is expected to depend on the number of years over which the estimation is made, we have used only countries for which the last survey was held after 2010, and included the survey year as control variable. Moreover, we have added robustness analysis in which we have restricted the sample to countries for which the survey was held after 2015 instead of 2010 (Appendix 4.II). The results are similar in terms of direction and significance.

4.3.2 Facebook indicators

To create subnational social media indicators, we used Facebook's advertisement platform. For marketing purposes, Facebook allows users with a Facebook account to search the number of active Facebook users in the last month disaggregated by various geographic and demographic characteristics (detailed information can be found at: https://developers.facebook.com/docs/marketing-apis). In this way, Facebook enables potential advertisers to use advanced targeting options by creating detailed user profiles, which include demographic characteristics like age, gender and geographic location. For example, if you want to reach females aged 20-30 in Benin, Facebook provides you the number of active accounts last month that fulfill these criteria. Earlier studies already used Facebook's advertisement platform for research purposes (Alexander et al., 2019; Stewart et al., 2019; Zagheni et al., 2017).

We used this possibility to create active user estimates for individuals aged 15-39 at the level of subnational regions in SSA countries. To obtain the percentage of Facebook coverage, we calculated the percentage using the number of Facebook users, compared to the number of individuals aged 15-39. The population sizes for these individuals were based on data from the most recent DHS and MICS surveys and data from the WDI, using the same estimation procedure as for the subnational CBR values (discussed above). First, we computed a differentiation factor for individuals aged 15-39 between the national population sizes in the DHS/MICS surveys and those in the 2020 WDI data. Thereafter, this differentiation factor was applied to the subnational population data in the DHS/MICS surveys to obtain the 2020 subnational population estimates.

For 19 subnational regions, the Facebook coverage exceeded the 100 percent. As Facebook is not transparent in how they calculate their audience sizes, we decided to set the maximum percentage of Facebook coverage at 100 percent. For robustness, we have also run the analysis without these subnational regions, with a maximum of 75 percent instead of 100 and with subnational Facebook deciles instead of percentages. As the results did not significantly differ in terms of direction and significance, we only present the results of the analysis that included the regions with a Facebook percentage up to 100 percent.

Besides the percentages of Facebook coverage in the subnational regions, we also computed the gender gap in Facebook coverage in the subnational regions. First, we calculated the Facebook coverage for males and females separately using the same technique as used to calculate the Facebook coverage for both genders. Second, we subtracted the subnational female Facebook coverage from the male Facebook coverage. As such, if the gender gap in Facebook coverage is zero this means equal coverage of males and females on Facebook. If the gender gap is positive, males are overrepresented, and if it is negative, females are overrepresented.

4.3.3 Control factors

Control factors in our study are the level of development, gender inequality, phone ownership, the percentage of females aged 15-49, and the age structure of fertile women in the regions. In more developed and wealthier regions, the digital access is better and the fertility rates are known to be lower (Pezzulo et al., 2021). As such, regions' development should be controlled for. Lower birth rates are also expected for regions with a stronger position of women, for regions with higher levels of mobile phone ownership, for regions with a lower percentage of females in their fertile years (ages 15-49), and for regions were the fertile women are on average older (Casterline, 2017; Rotondi et al., 2020).

The level of development is measured by the subnational Human Development Index (sHDI) (Smits & Permanyer, 2019). The sHDI shows within-country variation in human development across the globe, and is like the Human Development Index an indicator of a combination of three pillars: education, health and standard of living. The sHDI is measured with a value between 0 and 1 and a higher score denotes higher development. Gender inequality is measured by the subnational gender development index (sGDI), which is defined as the ratio between the female and male levels of sHDI in the region (Smits & Permanyer, 2020). A sGDI of 1 indicates perfect equality between females and males; a sGDI below 1 indicates that males have a higher level of human development and a sGDI above 1 indicates that

females have a higher level of human development. Phone coverage is measured by the percentage of households of which at least one of the members owns a phone in the region. The data are derived from the DHS, MICS and Afrobarometer surveys. Afrobarometer data on phone ownership for 2020-2021 (round 8) are used if they are more recent than the DHS or MICS data. This is the case for Benin, Cameroon, Cote d'Ivoire, Eswatini, Ethiopia, Gabon, Gambia, Ghana, Guinea, Kenya, Lesotho, Liberia, Mali, Mozambique (round 7, 2017), Namibia, Niger, Nigeria, Togo. The percentage of females aged 15-49, which is considered as females' fertile period, were calculated as the percentage of females aged 15-49 on the total population size of the region. The population structure of fertile females was included by calculating the mean age of females aged 15-49 in the region.

4.4 Methods

4.4.1 Exploratory spatial data analysis

We started the analyses with an exploratory spatial data analysis of the subnational crude birth rates, Facebook coverage and gender gaps in Facebook coverage. We examined whether spatial clustering exists among those three variables and to what extent there is overlap between clusters of birth rates and those of Facebook coverage. Clusters refer to agglomerations of bordering subnational regions which share similar characteristics. For example, adjacent regions that both show high (or low) crude birth rates or high (or low) Facebook coverage.

First, we assessed the degree of global spatial autocorrelation with the Global Moran's *I* statistic for the birth rates and Facebook coverage (Gittleman & Kot, 1990). Moran's *I* measures whether a spatial pattern exists. Global Moran's *I* values closer to 0 indicate a random spatial pattern, positive values indicate clustering and negative values dispersion. Second, we calculated the Gi* statistic based on local Getis-Ord Gi* hot spot analyses to examine whether and where clusters in birth rates and Facebook coverage were situated (Ord & Getis, 1995). The Gi* statistic identified areas where relatively high or low values of certain characteristics, in this case birth rates and Facebook coverage, tend to cluster together based on the neighbouring values. We compared the overlap in clusters of the birth rates and Facebook coverage and examined the characteristics of low and high birth rate clusters.

4.4.2 Regression analysis

To test whether the association between birth rates and Facebook coverage is still significant after the inclusion of subnational characteristics, we used regression

analysis. As the subnational regions are nested within countries, our data has a hierarchical structure. To control for this, two types of regression models were estimated: a mixed effects regression model which addresses the hierarchical structure by including a random intercept at the national level (Snijders & Bosker, 2011) and a fixed effects regression model in which the hierarchical structure is addressed by including fixed effects dummies at the country level (Allison, 2009). The fixed effects approach uses much more degrees of freedom than the multilevel approach, but has the advantage that all direct effects of (known and unknown) confounders at the national level are controlled for (Bell et al., 2019).

4.5 Results

4.5.1 Geographic patterns of Fertility Behavior and Social Media

Our analysis start with exploring the geographic pattern of subnational crude birth rates in 2020 (Figure 1). The mean subnational crude birth rates is 34 births per 1000 individuals in 2020, with a minimum of 17 births in North West, South-Africa and a maximum of 55 births in Lunda Sul, Angola. The mean Facebook coverage is 25%, with a minimum of 1% in Other Central, Malawi and a maximum of 100% in 19 regions (darkest color in Figure 1b). The gender gap in Facebook coverage is on average 8%, meaning that females are on average less covered than males. Females are relatively the best represented on Facebook in Eastern Cape, South Africa (8% more than males), and males show relatively the highest coverage in Maritime, Togo (59% more than females). Figure 4.1 shows substantial variation across subnational regions in the distribution of the crude birth rates, Facebook coverage and the gender gap in Facebook coverage. The descriptive statistics can be found in Table 4.1.

Figure 4.1: Geographic distributions of (a) crude birth rates 2020, (b) Facebook coverage 2021, and (c) the gender gap in Facebook coverage between males and females 2021

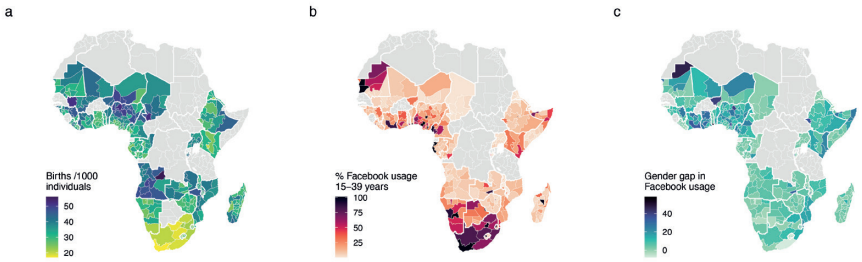


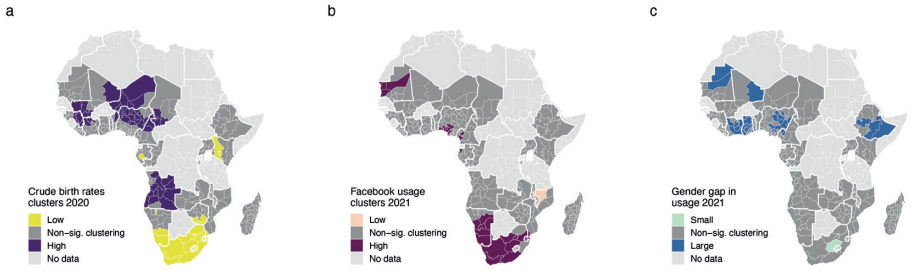
Table 4.1 Descriptive statistics: Means and percentages of variables included in this study

Variables	Mean, %	SD	Min.	Max.
Subnational Crude Birth Rates 2020 (dep. var.)	34.00	7.00	16.77	55.32
Subnational Characteristics				
% Facebook	24.56	28.42	0.55	100
Gender gap Facebook	8.18	9.88	-7.84	58.55
sHDI	0.53	0.09	0.27	0.75
sGDI	0.89	0.08	0.66	1.07
% Phone	77.74	18.85	22.50	99.02
% Females 15-49	23.37	5.15	16.58	50.57
Mean age females 15-49	28.62	0.98	25.53	31.39
Year of survey	2016	2.40	2011	2020

N regions=311, N countries=29

We tested whether the variability in these distributions showed a spatial pattern and found a Moran's *I* of 0.09 (SD=.005) for birth rates, of 0.07 (SD=.007) for Facebook coverage, and of 0.11 (SD=0.007) for gender gaps in Facebook coverage. This indicates that the variability in birth rates and Facebook coverage is not random and that clustering exists, whereby neighboring subnational regions tend to have similar characteristics and form spatial clusters. Figure 4.2 presents the Gi* statistics, which show where clusters in crude birth rates, Facebook coverage and gender gaps in Facebook are situated. Clusters with Gi* statistics below -1.96 are considered as significantly low (or small) clusters, and those with Gi* statistics above 1.96 are considered as significantly high (or large) clusters. The mean crude birth rate in low birth rates clusters was 25 births per 1000 individuals (SD=4.8), and in high birth rates clusters was 44 births per 1000 individuals (SD=4.5). The average Facebook coverage of low coverage clusters is 2.7 percent (SD=4.2) and of high coverage clusters is 48.4 percent (SD=20.3), the average gender gap in small gender gap clusters is 0.2% (SD=4.2) and in large gender gap clusters 21.0% (SD=11.5) (Appendix 4.II).

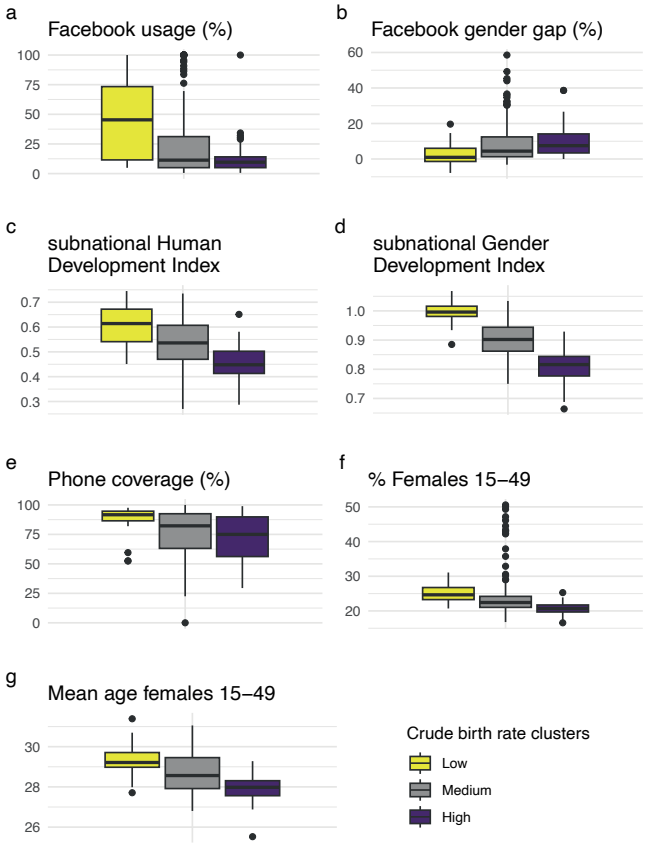
Figure 4.2 Spatial clusters of (a) crude birth rates 2020, (b) Facebook coverage 2021, and (c) the gender gap in Facebook coverage 2021



Visual inspection of figure 4.2 shows that clusters with low birth rates tend to overlap with clusters of high Facebook coverage, but that this is less so for gender gap clusters. Low birth rate clusters overlap with clusters characterized by high Facebook coverage in Southern SSA [Namibia, South Africa and Swaziland]. High birth rate clusters overlap with some clusters with large gender gaps, but not so much with clusters with high Facebook coverage. Here, we see that high birth clusters and large gender gaps in Facebook coverage predominantly overlap in some parts of Western SSA [Mali and Nigeria].

Further examination of the relationship between the clustering of subnational crude birth rates and Facebook indicators reveal consistent patterns. Figure 4.3a shows that the subnational Facebook coverage is on average higher in low birth rate clusters. However, there is no clear difference between the Facebook coverage of the non-significant and high birth rate clusters. These observations seem to confirm the visual overlap from figure 4.2, which showed that high Facebook coverage is associated with low crude birth rates, but a low Facebook coverage is not necessarily associated with high crude birth rates. Figure 4.3b shows that the gender gap in Facebook coverage is smaller in low birth rate clusters. The median of the gender gap is slightly above zero in low birth clusters, meaning that the Facebook coverage of males and females is (almost) equal in these clusters. On the other hand, the median of the gender gap in high birth rate clusters is 8, which shows that females are on around 8% less covered on Facebook than males in high birth rate clusters. As such, the gender gap in Facebook seems to be associated with birth rate clusters.

Figure 4.3: Boxplots showing means of the key variables for low and high birth rate clusters



Notes: Boxplots showing the minimum (lower end line), maximum (upper end line), median (middle line of box), and fist and third quartiles (lower and upper ends of box) of low, non-significant and high birth rates clusters of (a) the Facebook coverage 2021, (b) the gender gap in Facebook coverage 2021, (c) the subnational Human Development Index 2020, (d) the subnational Gender Development Index 2020, (e) the phone coverage, (f) the percentage of females aged 15-49, and (g) the mean age of females aged 15-49 for crude birth rates in low, non-significant and high birth rate clusters.

Similar patterns are observed between the crude birth rate clusters and the control factors: sHDI, sGDI, phone coverage, the percentage females aged 15-49 and the mean age of females aged 15-49. The sHDI, sGDI, phone coverage, percentage females aged 15-49 and the mean age of females aged 15-49 are higher in low birth rate clusters. This shows that low birth rate clusters have on average higher levels of development, levels of gender development, percentages of phone coverage and percentages of females aged 15-49.

4.5.2 Regression results

Table 4.2 shows the results of the regression models which test the association between birth rates and Facebook coverage. The mixed effects model and the fixed effects model show comparable results in terms of coefficients, standard errors and significance.

Table 4.2. Regression estimates of the determinants of subnational Crude Birth Rates (2020) for the multilevel linear regression (MLR) and Fixed Effects regression models.

	MLR I	**FE I**
% Facebook	-0.030 * (0.014)	-0.035 * (0.015)
Gender gap Facebook	0.062 * (0.031)	0.067 * (0.032)
sHDI	-21.852 *** (6.357)	-24.025 ** (7.711)
sGDI	-31.128 *** (7.509)	-33.947 ** (10.611)
% Phone	-0.064 * (0.025)	-0.073 * (0.03)
% Females 15-49	0.173 (0.11)	0.364 * (0.162)
Mean age females 15-49	-0.934 ** (0.35)	-0.67 (0.379)
Year	0.071 (0.265)	
Country dummies included		Yes
Constant	33.529 *** (0.636)	41.317 *** (1.396)
Country-level variance	9.62 (3.10)	
Adjusted R^2		0.657

Notes: * $p<.05$, ** $p<.01$, *** $p<.001$; Independent variables are centered around their mean. Standard errors are shown in parentheses. Analyses are run for 311 subnational regions in 29 countries.

Our analysis reveal significant associations of Facebook coverage and gender gaps in Facebook coverage with subnational birth rates. Facebook coverage is negatively associated with birth rates, which means that fertility is lower in areas with a higher Facebook coverage. This finding is in line with our first hypothesis.

Regarding the gender gap in social media coverage, we find a positive association between the gender gap in Facebook coverage and birth rates. Hence in areas where females are relatively underrepresented on social media fertility levels are higher. This finding is line with our second hypothesis. We expected that females who have relatively less access to Facebook are less empowered and consequently, are less likely to get in contact with globalized ideas and to change their behavior according to these views.

All associations are controlled for the sHDI, sGDI, subnational phone coverage, percentage of females aged 15-49, mean age of females aged 15-49, and survey year. The association between birth rates and the sHDI is significantly negative, meaning than when subnational regions have a higher level of development, crude birth rates are lower. Birth rates are significantly higher in regions with higher gender inequality in favor of males. Subnational phone coverage shows a significant negative association with subnational birth rates, which means that a higher phone coverage is related to lower birth rates. The percentage of females aged 15-49 is positively associated with birth rates and the mean age of women in this age group negatively. Hence fertility is higher if fertile females comprise a larger share of the population and if the mean age of them is lower. The year of the survey is not significantly related to the birth rates. All these findings for the control variables are in line with our expectations and with earlier findings (Atake & Gnakou Ali, 2019; Billari et al., 2020; Casterline, 2017; Pezzulo et al., 2021; Rotondi et al., 2020).

4.6 Discussion

Where previous studies have firmly established a connection between fertility behavior and traditional media, like television and radio (Ferrara et al., 2012; Inkeles & Smith, 1974; Lerner, 1958; Westoff et al., 1997), this study examines the link between fertility behavior and social media. Based on data for 311 subnational regions in 29 countries in SSA, we tested two hypotheses on the relationship between social media and fertility outcomes: the idea that access to social media is associated to lower fertility rates (hypothesis 1) and the idea that gender equality in social media use is associated with lower fertility rates (hypothesis 2). Our findings are in line with the predictions of these hypotheses. The descriptive analysis identified spatial overlap between low birth rate clusters and high social media coverage clusters, and between high birth rate clusters and clusters with a large gender gap in social media. The regression analyses revealed that higher social media coverage and smaller gender gaps in social media coverage are significantly linked to lower birth rates. We therefore conclude that on the basis of our findings the hypothesis cannot be rejected and that our study provides new empirical evidence regarding the association between social media and fertility decline.

This study is important, because it is the first that shows the existence of a substantial relationship between Facebook coverage and fertility at the level of subnational regions in SSA. Although some studies already reported effects of mobile phone ownership and internet access on fertility behavior (Billari et al., 2020;

Toffolutti et al., 2020), we went a step further by identifying similar effects of social media exposure and in particular of access to Facebook. One possible explanation for this finding is that social media enable their users to form (online) communities and to share ideas among them. Interactions between individuals on social media rely less on physical proximity, and as such individuals are likely to encounter more globalized ideas and norms, and to change their behavior accordingly. As the discussion of new ideas and norms is a central concept within diffusion theories (Coale & Watkins, 1986; Westoff & Koffman, 2011), the findings of this study could be well interpret within this line of theory.

Our results regarding the gender gap in social media coverage reconfirm the finding of earlier studies that women are underrepresented in the online population (Fatehkia et al., 2018; Kashyap et al., 2020; L. Robinson et al., 2015). That better representation is associated with lower fertility outcomes is also in line with earlier studies who found that females adopting new and social media tend to have more modern and liberal values (Abubakar & Dasuki, 2018; Varriale et al., 2022). As such, our study also points towards negative consequences of large gender inequalities in media access.

A possible limitation of the present study is that there are potential reliability issues with the social media data. The finding of a Facebook coverage above 100 percent for 19 subnational regions is exemplary for this. It might be that individuals in those regions have more than one Facebook account or that there are a substantial number of fake accounts. According to (Meta, 2019), on average about 5% of all Facebook accounts are fake accounts. However, it can also be the result of errors in Facebook's target audience generator. The consequences could be twofold. If overestimation is systematic, it will not substantially influence the direction and significance of the regression estimates. If overestimation is random, this leads to the presence of measurement error in our data, which most likely leads to an underestimation of the real effects (Hutcheon et al., 2010). A second potential limitation is the possibility of selection bias in our data. It is possible that the regions with a higher social media coverage are also the regions with more globalized values and lower fertility rates. If so, this could lead to an overestimation of the association between social media access and fertility. More research is required to examine this possibility. A third limitation is that we have estimated the subnational crude birth rates by applying subnational variation derived from earlier surveys to national data derived from the World Bank. This approach is based on the assumption that subnational variation is relatively stable over time. To be sure, we added the year of the survey as an additional control factor and performed

a robustness test by repeating our analysis with a sample of only countries with surveys from 2015. Findings turned out to be substantially the same as in the main analysis.

In sum, this study uses an innovative approach to provide new evidence that social media coverage and gender gaps in coverage are significantly related to SSA fertility outcomes. This might mean that the social media revolution which currently spreads across the subcontinent might profoundly affect its fertility transition. Compared with earlier studies, our research constitutes a major improvement because it uses – for the first time – social media coverage data at subnational level as potential ingredient of fertility decline in the SSA region.

Chapter 5: The effect of mobile internet on child marriage and teen pregnancy in sub-Saharan Africa[12]

12. This chapter is based on joint work with Karlijn Morsink and Natascha Wagner and has received a revise and resubmit decision from the Journal of Development Economics that is due in April 2026.

The study on which this chapter is based was presented at the 2024 Dutch Development Economics Workshop in Wageningen, the 2024 Development Economics workshop in Tuscany 2024, the Economics Idea Lab at Radboud University in 2024, the 2025 German Development Economics Conference, and at the International Institute of Social Studies of Erasmus University Rotterdam in 2025.

Child marriage and teen pregnancy remain prevalent in sub-Saharan Africa. This paper examines the impact of mobile broadband internet (3G) expansion on these outcomes. We combine geo-referenced data on 3G coverage with individual-level marriage and fertility data from over 500,000 women-year observations across 20 sub-Saharan African countries between 2011 and 2022. Using two-way fixed effects, event studies, a stacked difference-in-differences, and an instrumental variables strategy, we find that adolescent girls exposed to 3G are between 1 and 6 percentage points less likely to marry and approximately 1 percentage point less likely to give birth, with no significant effects for older women. The effects are strongest among wealthier and more educated adolescents girls. Exploratory analyses suggest that increased access to information about family planning, shifting gender attitudes, and behavioral changes serve as mechanisms. Our findings imply that the expansion of 3G may be more effective than many conventional family planning interventions.

5.1 Introduction

Sub-Saharan Africa (SSA) has the highest rates of child marriage and teen pregnancy globally, with both phenomena linked to adverse health and economic outcomes (Chari et al., 2017; Field & Ambrus, 2008; Ganchimeg et al., 2014; Jensen & Thornton, 2003; WHO, 2024). According to the Child Marriage Data Portal (2024), one in five women worldwide were married before age 18, with prevalence highest in SSA. Similarly, UNICEF (2024) estimates that in 2022, 13% of women in SSA gave birth before turning 18. While child marriage has declined in Central and Southern Asia, progress in SSA has stalled over the past decade. Policy interventions aimed at reducing early marriage and fertility have shown limited effectiveness (Feyissa et al., 2023), and strategies that succeeded in Asia – such as large-scale family planning programs – have been less effective in Africa, where religious norms and lower acceptance of fertility control play a role (Caldwell & Caldwell, 2002). Emerging evidence suggests that access to media and information and communication technology (ICT) can influence fertility and marriage decisions (Billari et al., 2020; Bönisch & Hyll, 2023; Cassidy et al., 2024; Dewi et al., 2018; Ferrara et al., 2012; Guldi & Herbst, 2017; Kearney & Levine, 2015). This paper examines whether the rapid expansion of mobile broadband internet (3G) in SSA since 2011 – a key driver of social media access – has affected child marriage and teen pregnancy.

Our study draws on marriage information and records of birth histories from the Demographic and Health Surveys (DHS) and the geo-referenced Multiple Indicator Cluster Surveys (MICS). Our main dataset consists of 528,926 women-year observations gathered from 158,046 women and 20 DHS and 6 MICS surveys conducted between 2014 and 2022 covering 20 sub-Saharan countries. We combine the survey data with the geo-referenced mobile phone coverage data from the GSM Association (GSMA). We identify whether mobile broadband internet, specifically 3G or higher, is available in the woman's location of residence in the year of marriage and in the year before a child is born[13]. We then test if the expansion of 3G coverage influences child marriage and teen pregnancy using a two-way fixed effects model, a stacked diff-in-diff, an event study and an instrumental variables approach. In the TWFE and DiD models we include a variety of woman-level characteristics, time varying locational controls, location-of-residence (cluster) dummies, country-year fixed effects, and woman-fixed effects. In the IV model, we control for geographical

13. 3G mobile broadband, as opposed to 2G mobile broadband, introduced significantly higher-speed mobile internet, improved voice quality, and enhanced multimedia capabilities, such as use of social media.

and infrastructural characteristics of clusters, as the lightning instrument is only conditionally random.

Across all specifications, we find that 3G coverage significantly reduces child marriage and teen pregnancy, and effect sizes and significance levels are robust to choice of specification. Depending on the model, adolescent girls are between approximately 1 to 6 percentage points less likely to get married, and approximately 1 percentage point less likely to give birth when covered by 3G, while older women are unaffected. The economic magnitudes of these effects are substantial, as they imply an 8% reduction in child marriage and a 7% reduction in teen fertility, relative to control means of 24% and 20%, respectively. These effects are similar in magnitude to effect sizes of broadband internet and TV in Brazil, Europe and the United States; and are promising given the comparatively low success of family programs in SSA.

Heterogeneity analyses show that the effects on child marriage and teen pregnancy are particularly strong for wealthier and more educated women, suggesting that these women are likely the early adopters of the new technology. We perform complementary exploratory analyses to explore mechanisms, which suggest that knowledge about family planning, attitudes towards women and children and behavioral change, including the use of contraception, are likely to be drivers of the observed reductions in child marriage and teen pregnancy.

The identifying assumption underlying the TWFE model is that changes in 3G coverage in a given location and year are – conditional on control variables – as good as random. The main threat to the validity of the TWFE estimator is that it potentially suffers from bias under dynamic treatment effects (Borusyak et al., 2024; Callaway & Sant'Anna, 2021; Goodman-Bacon, 2021; Sun & Abraham, 2021). To address this potential bias we employ a stacked difference-in-difference (DiD) analysis (Goodman-Bacon, 2021; Wing et al., 2024), an event study (Sun & Abraham, 2021), and we use the imputation approach (Borusyak et al., 2024). TWFE, DiD and event study methods are valid estimators conditional on the lack of differential pre-trends. While we show that there exist no differential pre-trends for the fertility analyses, we cannot confidently rule out differential pre-trends in the child marriage analyses. Therefore, we also implement an instrumental model using time-interacted lightning strikes as instrument (Flückiger & Ludwig, 2023; Manacorda & Tesei, 2020). For the teen pregnancy results, where we observe no pre-trends, all estimates fall within the confidence intervals of the TWFE estimates. For the child marriage results, where we cannot confidently rule out pre-trends,

the estimates from the DiD, event study, and imputation methods fall within the confidence intervals of the TWFE estimates, while the IV estimates differ slightly, but are still of similar magnitude.

We connect to a growing literature on the role of media in shaping gender norms, marriage, and fertility decisions. Media can modify important drivers of marriage and fertility through social learning and social influence or both (Balbo & Barban, 2014; Bernardi et al., 2007; Bongaarts & Watkins, 1996). Already in 1974, Inkeles and Smith (1985) showed that watching television and listening to radio installs different norms and values like limiting family size. Several television series, have shown to reduce (teen) fertility in both low- and high-fertility settings, such as Brazil and the US (Ferrara et al., 2012; Kearney & Levine, 2015). In a similar vein, the expansion of private television broadcasts had fertility-reducing effects in Germany and Indonesia (Bönisch & Hyll, 2023; Dewi et al., 2018). Moreover, exposure to cable television empowered women in India by increasing school enrollment, birth spacing, self-reported autonomy and by decreasing acceptability of domestic violence (Jensen & Oster, 2009). Exposure to edutainment via mobile cinema reduced child marriages in Pakistan (Cassidy et al., 2024). Likewise, the internet has been shown to have fertility-reducing effects for teenagers in the US, women's phone ownership has reduced fertility in Malawi, and mobile broadband internet reduced the prevalence of traditional social norms, including lower marriage and fertility rates in Jordan (Billari et al., 2020; Guldi & Herbst, 2017; Viollaz & Winkler, 2022).

More specifically, we contribute to a literature on the effects of mobile phones and 3G on development, which has so far been documented for political mobilization (Manacorda & Tesei, 2020), increases in women's political empowerment in Africa (Nkoa et al., 2023), reduction of corruption, improvement of confidence in the government (Guriev et al., 2021), increased labor force participation (Chiplunkar & Goldberg, 2022; Viollaz & Winkler, 2022), increased welfare among Indian fishermen (Jensen, 2007), and reduced poverty (Bahia et al., 2024). Mobile phones have also been documented to reduce fertility, lower gender inequality, increase contraceptive uptake, and lower maternal and child mortality, with the payoffs being largest in disadvantaged areas (Billari et al., 2020; Flückiger & Ludwig, 2023; Rotondi et al., 2020; Toffolutti et al., 2020).

This research also contributes to the role of multidimensional inequalities surrounding ICT, often referred to as the digital divide (DiMaggio, Hargittai, Neuman, et al., 2001; Norris, 2001) and how it relates to who benefits most from

new technology. First, there is a gender divide (Alozie et al., 2018; UNICEF, 2023). Young African girls have lower access to mobile phones than young boys, and when girls have a mobile phone in rural areas, they are sometimes even considered as 'prostitutes' (Porter et al., 2020). African girls, compared to African boys, are less likely to use social media like Facebook (Fatehkia et al., 2018; Wildeman et al., 2023). Second, there is an urban-rural divide. Phone ownership is substantially higher in urban compared to rural areas, independent of both gender and age (Lembani et al., 2020; Porter et al., 2020). By studying heterogeneities in our sample we further contribute to quantifying the digital divide.

The remainder of the paper is structured as follows. Section 2 provides some background on child marriage, teen pregnancy, and 3G coverage in SSA, and describes potential relationships. Section 3 presents the data and construction of the estimation sample, and section 4 introduces our estimation strategy. The results including heterogeneity analyses and mechanisms are discussed in section 5 and section 6 concludes.

5.2 Background

5.2.1 Child marriage, teen pregnancy and 3G

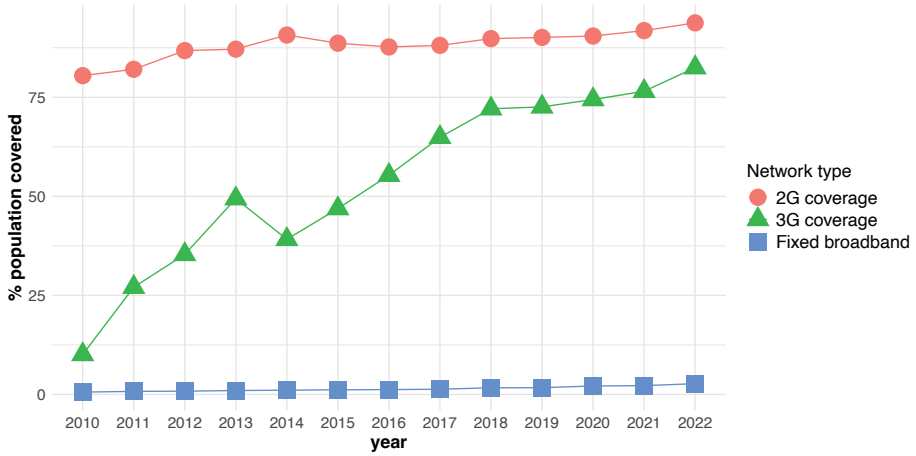
Marriage is an important institution in the African culture as it signals the beginning of a woman's reproductive life. There are four traditional characteristics of marriage in SSA: (1) marriage was virtually universal for men and women in the 1950s and 1960s; (2) women are often considered marriageable when they reach puberty, resulting in early ages of marriage for women (3) typically there exist large age differences between husbands and wives; (4) polygyny is common (Garenne, 2022). In the 20th century, changes in marriage patterns occurred, particularly after independence, i.e., since 1960: age of first marriage for women started to increase and marriage rates for both men and women started to decrease (Shapiro & Gebreselassie, 2014). Improvements in women's education and income are considered to be the key drivers behind the rising age of marriage in 32 SSA countries (Garenne, 2004). However, until today, SSA has the highest rates of child marriage worldwide, and 9 out of 10 countries with the highest child marriage rates are situated in SSA (UNICEF, 2015). The negative consequences of early marriage for many development outcomes such as education, health of mothers and children, and intimate partner violence and labor market participation for women are well documented (e.g. Chari et al., 2017; Field & Ambrus, 2008; Hicks & Hicks, 2019; Jensen & Thornton, 2003).

Giving birth to a child is another major event in a woman's life. Motherhood is considered a highly desirable role for women in most African societies (Ahonsi et al., 2019; Schaffnit et al., 2019). Moreover, in SSA, women give birth to their first child early compared to other low- and middle-income countries (Kassa et al., 2018; UNICEF, 2015). As such, it might not be surprising that SSA has the highest teen pregnancy rates worldwide (Kassa et al., 2018; WHO, 2024; Yakubu & Salisu, 2018). In 2023, the adolescent birth rate for girls aged 10–14 years was 4.4 per 1000 women. For girls aged 15-19 the birth rate was 97.9 per 1000 women, compared to a global birth rate of 41.3 per 1000 women. Teen pregnancy has been linked to serious health issues for the mother and her child (e.g., low birth weight/premature birth, anemia, high blood pressure/pregnancy-induced hypertension, higher rate of infant mortality), school dropout, poverty, and domestic violence (Branson & Byker, 2018; UNICEF, 2015).

Teen pregnancy and child marriage are closely related and mutually reinforce each other, especially in SSA (Zgambo, 2022). Exemplary of this interrelation is existing evidence that delaying or preventing child marriage can help reduce teen pregnancy (UNFPA, 2012). In turn, teen pregnancy may be the gateway to child marriage (Agege et al., 2018), yet it can also be a consequence of child marriage (Menon et al., 2018). Once married, girls in SSA are expected to deliver a child without much delay, regardless of age (Ahonsi et al., 2019; Schaffnit et al., 2019).

Whereas the tradition of early marriage and fertility has been stable for many generations, 3G has developed very recently and rapidly. Figure 5.1 shows that 3G coverage in SSA grew from less than 10% in 2010 to more than 80% in 2022 (International Telecom Union, 2024). The 3G mobile network is a first-generation network that allows its users to freely browse the internet and use social media applications. The 3G signal affects internet usage on the extensive margin – by influencing the ability to go online – and on the intensive margin – by influencing the amount of time spent online – and qualitatively – by influencing what people do online (Guriev et al., 2021). In contrast to this sharp increase in 3G availability, fixed broadband internet access (ADSL or fiber-optic cables) hardly increased. As such, it is logical that the majority of Africans access the internet through their mobile phones (Bahia et al., 2024). They do so either on their own mobile phone or on the mobile phones of others, as African societies still incorporate strong sharing norms (Porter et al., 2015). This unique situation of 3G expansion in SSA sets the stage for our analysis.

Figure 5.1 Development of 2G, 3G and fixed broadband in sub-Saharan Africa



Notes: The figure shows the mean % of the population covered by 2G, 3G and the fixed broadband connections in sub-Saharan Africa. Source: ITU Digital Development: ICT Indicators Database (International Telecom Union, 2024)

5.2.3 Pathways from 3G to child marriage and teen pregnancy

We hypothesize that the expansion of the 3G network may influence child marriage and teen pregnancy through the provision of information, that may in turn change attitudes, preferences, beliefs and behavior[14]. First, 3G coverage may reduce the cost of searching for information (Aker & Mbiti, 2010), thus improving the quality of adolescent girls' decisions. The 3G network allows its users to obtain information immediately, without having to wait for radio broadcasts or newspapers. Moreover, adolescent girls can take an active role in the search process by asking questions and looking for responses from multiple sources. For example, a teenager can go online, search, and identify where she can get contraceptives or where she can see a doctor. On the other hand, they may accidentally – by browsing the web – come across information on social media, in chat groups, or encounter advertisements that could change her behavior.

Information has proven of high relevance for fertility and marriage behavior. For example, Bau et al. (2024) show that providing information that modern contraceptives do not cause infertility at an older age almost tripled contraceptive

14. Appendix D A.5.5 shows the conceptual model underlying how 3G coverage might influence child marriage and teen pregnancy through the information and changes in attitudes and preferences pathways. We set out from the cooperative bargaining model of the household and consider the availability of 3G as an alteration to the threat point.

use compared to only giving a conditional cash transfer. Other examples of the value of information include (1) improved health knowledge due to mobile phone coverage which is likely to be an important explanation for reduced child mortality in SSA (Flückiger & Ludwig, 2023), and (2) decreased information search costs which are likely to explain how mobile phone ownership reduced fertility in Malawi (Billari et al., 2020).

Information may also change attitudes, beliefs and preferences toward what is acceptable and preferable behavior. One important feature of 3G is the ease of accessing social media, such as WhatsApp and Telegram. It may also allow users to engage online with other users in social networks (Rainie & Wellman, 2012). Online communities enable individuals to discuss ideas and information without physical proximity, resulting in contact with persons living in areas where other gender norms, fertility, and marriage patterns prevail. Moreover, the online environment allows individuals to challenge the cultural norms of their communities and to ask questions they would not dare to ask face-to-face without any fear of repercussions or real-life social costs, potentially allowing for a faster trickle-down of different ways of life (Rainie & Wellman, 2012).

Other studies have already reported the importance of changes in attitudes and preferences triggered by the availability of ICT and information. For example, soap operas have influenced women's fertility by changing family size preferences (Ferrara et al., 2012). Rotondi et al. (2020) and Arias (2019) show that improved access to information can change social norms towards gender equality and empower women to make more independent decisions. Abubakar and Dasuki (2018) show that WhatsApp, as one form of social media, enabled women within a community in Nigeria to push the barriers created by social norms dictating gender roles. Contact with other women in these WhatsApp group chats increased their autonomy and allowed them to voice their personal opinions.

3G coverage may also effect child marriage and fertility through spillovers from those with access to 3G to adolescent girls. Other studies found that having access to 2G increased women's labor market participation and income (Bahia et al., 2024). Access to mobile money, which is facilitated by mobile coverage, may also serve as protection against income shocks, and as such serve as protection against child marriage in certain contexts (Corno et al., 2020; Corno & Voena, 2016). Finally, 3G coverage can also increase screen time, increasing the opportunity cost of time and thus making dating or sexual intercourse less likely.

5.3 Data

Information on women's marriage and fertility behavior along with other women and household characteristics are taken from the Demographic and Health Surveys (DHS)[15] and the Multiple Indicator Cluster Surveys (MICS)[16]. These surveys are nationally representative cross-sectional household surveys that collect data on women's marriage and full birth histories[17]. Our main dataset uses information from 20 DHS and 6 MICS surveys conducted between 2014 and 2022 and covers 20 sub-Saharan African countries, some of which are surveyed more than once (see Table A2 for a detailed list of all countries and surveys included).

As child marriage is typically defined as marriage before the age of 18, we select women between the ages 10 to 17 for our main models on child marriage. We construct the variable marriage as a dummy variable equal to 1 if a woman *i* in year *t* got married, and zero otherwise. In case a woman got married in a particular year, all subsequent year observations for that woman are coded as missing, as we assume that (child) marriage happens once, and a child married woman is unlikely to have divorced, and re-married again before she passes the age 18 threshold. For the additional analyses, to assess whether 3G also affects women's chances of getting married at older ages, we change the age selection of our sample in seven age groups [10-14; 15-19; 20-24; 25-29; 30-34; 35-39; 40-49]. The latter group spans 10 years to make sure we observe enough marriages as most women are typically married before the age of 40.

We define teen pregnancy as giving birth to a child before the age of 20. Therefore we use a dataset of women between ages 10 and 19 for our main models on teen pregnancy. We construct the variable "got child" as a dummy variable equal to 1 if a woman *i* in year *t* gave birth, and zero otherwise. Unlike the marriage data, a woman stays in the data after she gave birth to a child as she can give birth to a subsequent child even during her teens. We also test whether 3G affects women's

15. www.dhsprogram.com
16. mics.unicef.org
17. The sampling strategies of these surveys – to create a nationally representative, randomly selected sample of households – are to first randomly select a subset of non-overlapping geographical areas, called "clusters," which typically comprise communities, villages, towns, cities, or neighborhoods. In each of the selected clusters, all households are listed, and a random subset of 25-30 households are selected to participate in the surveys. Household members are first asked for basic household information, including a household roster before women aged 15 to 49 years are interviewed.

probability of giving birth at older ages using the same age groups as described above for the child marriage analysis.

To measure 3G coverage, we use mobile coverage explorer data, which provides yearly representations of the area covered by mobile networks around the world. It has been created from submissions made directly to Collins Bartholomew[18] or the GSM Association (GSMA), which is an association that represents the interests of the mobile phone industry worldwide. This data are supplemented with coverage data created from OpenCellID, the world's largest open crowd information database of cell towers since 2019. We use digital maps of global 3G network coverage from 2011 to 2021 as 2011 marks the year when GSMA data on 3G for Africa became first available. As we are interested in the effect of having access to mobile broadband internet we consider not only 3G access but also include further improved access such as 4G. For each of the available country years, we identify the clusters with broadband internet coverage and refer to it as 3G coverage because 3G is by far the most prevalent form of available broadband internet.

We combine 3G coverage data with the DHS and MICS surveys based on the geocoded location of the residence of the woman. The main explanatory variable is a dummy variable that captures whether 3G coverage is available at the cluster location in a given year or not.

For our child marriage models, we have 490,946 women-year observations from 133,896 women. For our fertility models, we have 528,926 women-year observations gathered from 158,046 women. The former dataset is smaller because women that got married are coded as missing for subsequent years and the age restriction is more narrow. On average, a woman is observed for 3.3 years in the child marriage data, and for 3.7 years in the teen pregnancy data. Note that there is a seeming disconnect between the sample period of the household surveys and the 3G data as the earliest household survey data are from 2014. Yet, the final sample spans the period 2011 to 2021 since we make use of the DHS and MICS marriage and birth histories allowing us to backwardly expand into the years 2013, 2012, and 2011.

Table 5.1 shows the descriptive statistics for our main outcome and control variables. The first row shows that 19.6% of the teenagers in our sample gave birth to at least one child within the years they are included in our study. The yearly incidence rate of giving birth is 6.7%. The total incidence rate of child marriage in the sample is 23.9%, and the yearly incidence rate of marrying is 3.7%.

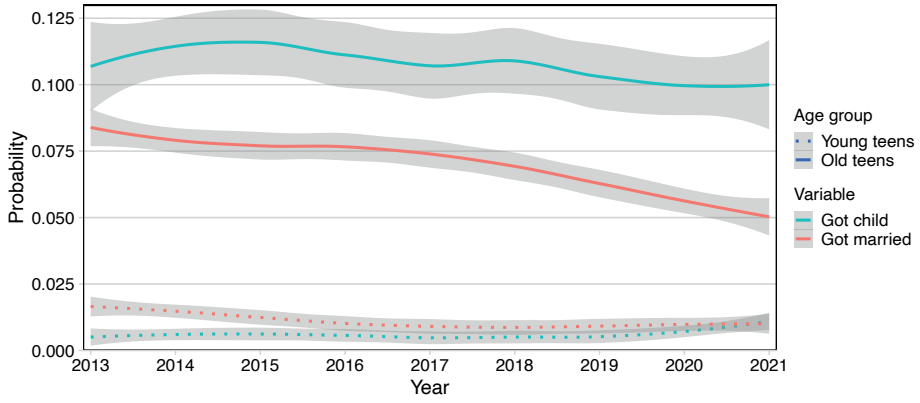
18. https://www.collinsbartholomew.com/mobile-coverage-maps/

Table 5.1: Descriptive statistics of the main datasets

	Child marriage sample	Teen pregnancy sample
Mean prob. got married overall	0.239	
Mean prob. got married yearly	0.037	
Mean prob. got child overall		0.196
Mean prob. got child yearly		0.067
Mean 3G coverage	0.288	0.256
Minimum age	10	10
Maximum age	17	19
N women-year observations	490,946	528,926
N woman	133,896	158,046
Mean # obs. per woman	3.667	3.347

Over time we see a slight decrease in child marriage starting from 2017, especially for older adolescents aged 15 to 17 years (Figure 5.2). The decline in teen pregnancy for older adolescents (15 to 19 years) is less pronounced as also indicated by the wider confidence bands. For younger adolescents, aged 10 to 14 years, the marriage and pregnancy incidence rates are lower, yet remarkably stable over time.

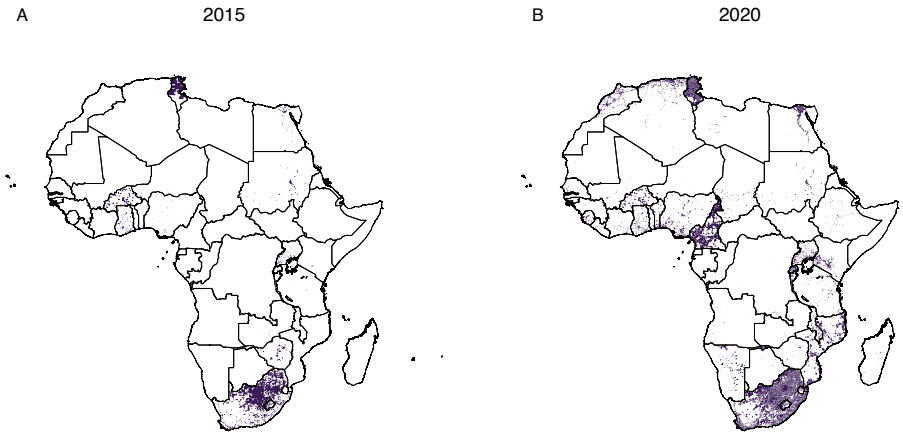
Figure 5.2 Trend in child marriage and teen pregnancy between 2013 and 2021



Notes: The figure shows the mean probability of giving birth and getting married for young teens (aged 10-14) in dotted lines and old teens (aged 15-19 for giving birth and aged 15-17 for getting married) in solid lines, for the sample of adolescent girls used in this paper. 2011 and 2012 are not included in this graph as only have data on two countries in those years. As such, it is not representative of the trend in the overall sample.

Average mobile internet coverage is low at around 25%. However, this average masks considerable spatial and temporal variation as shown in Figure 5.3 and Appendix D Figure A5.1. The first available GSMA data on 3G coverage in Africa is from 2011, but only for a limited number of countries. Over the years, the number of countries with 3G data increased as well as the coverage rate within countries. We make use of the temporal and geographical variation in access to 3G to identify whether this variation influences child marriage and teen pregnancy.

Figure 5.3: Expansion of 3G coverage between 2015 and 2020



Notes: Panel A shows the availability of 3G coverage in Africa in 2015. Panel B shows this availability for 2020, as reported by GSMA.

5.4 Empirical framework

5.4.1 Two-Way Fixed Effects (TWFE)

The source of variation for the two-way fixed effects (TWFE) analysis comes from the staggered roll-out of the 3G network. This staggered roll-out generates geographic and temporal variation in 3G availability and, consequently, in child marriage and teen pregnancy. First, we estimate the effect of 3G coverage by comparing marriage and fertility outcomes of adolescent girls in places where 3G was and was not (control group) rolled out at two moments in time, namely before and after the arrival of the 3G signal. As such, the basis for our empirical analysis is the following linear regression model:

$$Outcome_{icjt} = \beta_0 + \beta_1 3Gcov_{cjt} + \beta_2 X'_{i\mu ct} + \beta_3 C'_{\mu ct} + \rho_c + \tau_{jt} + \zeta_i + \eta_{icjt} \quad (5.1)$$

where $Outcome_{icjt}$ is a dummy indicating whether an adolescent girl *i* living in cluster *c* in country *j* at time *t* got married, our first outcome of interest, or gave birth, our second outcome of interest. The main regressor $3Gcov_{cjt}$ is a dummy variable that takes the value 1 if the cluster where the adolescent girl lives has 3G coverage, as opposed to no 3G coverage at time *t*. For our fertility models, we use lagged variables (t-1) because, from the time of conception, it takes 9 months before a child is born and we are interested in the effect of 3G around conception and not birth.

$X'_{i\mu ct}$ is a matrix of individual control variables. In its most demanding version, this set of controls consists of fixed effects for age and the availability of 2G coverage. $C'_{\mu ct}$ is a vector of time-varying cluster controls at time *t* (and $t-1$ for our fertility models) as well as time-invariant cluster controls. These include the frequency of negative and positive rainfall shocks, mean yearly rainfall, night-time lights, minimum and maximum yearly temperature, interactions between a linear time trend and elevation, distance to the coast, distance to a main river, and distance to a nearby city. Appendix D table A5.1 presents the data sources used to create these cluster controls, along with a short description of the content of these locational controls. Cluster fixed effects ρ_c absorb any time-invariant local characteristics, and the country x year fixed effects τ_{jt} account for country specific annual shocks, like country-wide economic recessions. Additionally, we control for any time invariant differences in woman characteristics by including woman-fixed effects ζ_i. Standard errors are clustered at the cluster level to account for the correlation of the error term across individuals within clusters.

The main threat to the validity of our empirical approach is that the TWFE estimator potentially suffers from bias under dynamic treatment effects, especially if treatment areas are not only compared to control areas, but also to areas treated earlier, or when there are heterogeneous treatment effects across groups due to differences in the timing of treatment (Borusyak et al., 2024; Callaway & Sant'Anna, 2021; Goodman-Bacon, 2021; Sun & Abraham, 2021). To provide additional support for the validity of our estimates, we employ four complementary approaches: (i) a stacked difference-in-difference (DID) analysis (Wing et al., 2024), (ii) an event study (Sun & Abraham, 2021), (iii) the imputation approach (Borusyak et al., 2024), and (iv) an instrumental variables (IV) model. In what follows we briefly detail the rationale for these models. Despite the acknowledged shortcomings of the TWFE estimator, we present the

TWFE results as our main findings as they are perceived as easily comprehensible and intuitive. More importantly, for the data at hand, the complementary empirical approaches yield virtually identical results. We also challenge the robustness of our estimates to controlling for a wide range of locational controls, like rainfall shocks and night-time lights as proxies for local development. We also estimate a series of models in which we include region x time trends, city x time trends, select different sources of 3G coverage and leave some countries out of the sample (see Appendix D A.5.2.5 for details).

5.4.2 Stacked DID and event study

First, we conduct a staggered, stacked (DID) analysis in which forbidden comparisons between regions that are both already treated but at different times are addressed (Wing et al., 2024). This method overcomes the limitations of the TWFE model by explicitly estimating separate treatment effects for each woman-year observation in our dataset, providing a more granular view of treatment impacts of 3G on child marriage and teen pregnancy, respectively. The regression is weighted by the cohort-specific synthetic unit weights applying Wing et al. (2024) weighting scheme. More details on the specification of this estimator are provided in Appendix D A.5.2.1.

Second, we use the event study model by Sun and Abraham (2021) to account for potential dynamic treatment effects. The model is built on the DID framework but allows for multiple time periods and variation in the timing of treatment. In addition, this method avoids the bias introduced by assuming homogeneous treatment effects, thus providing more flexible and accurate estimates of the causal impact of the impact of 3G coverage. For more details, see Appendix D A.5.2.2.

Third, to further refine our estimates, we add the imputation method proposed by Borusyak et al. (2024). This method provides an alternative for estimating the effect of the staggered 3G introduction, allowing for arbitrary heterogeneity in treatment effects and dynamics over time. The imputation method estimates a model on the not (yet) treated women and extrapolates this model to the women who gain 3G coverage. Appendix D A.5.2.3 provides more details.

All three models come from the DID class and might not be fully able to address any remaining endogeneity. Therefore, we complement these models with an IV model.

5.4.3 Instrumental variable approach

To address concerns about differential pre-trends, specifically for the child marriage analysis, and remaining concerns about endogeneity, we also employ an IV model.

The IV model exploits the fact that frequent lightning strikes – which have been shown to be conditionally random when controlling for a range of geographical and infrastructural characteristics of regions – delay the adoption of mobile phone technologies as established by in earlier studies Flückiger and Ludwig (2023); Guriev et al. (2021); Manacorda and Tesei (2020); Rotondi et al. (2020). Electrostatic discharges during storms damage mobile phone infrastructure and, in particular, antennas on the ground that transmit the signal in their vicinity, thus negatively affecting connectivity. The high costs of lightning conductors and other surge protection and weaker connectivity makes investment in 3G less profitable in areas with relatively more lightning strikes. Furthermore, poor connectivity makes the adoption of mobile phone technology less appealing. As a result, one would expect slower adoption of mobile phone technology in areas with high frequency of lightning strikes that fades out as more and more regions are covered. Therefore, we use local lightning frequency interacted with a time trend together with the squared term of this interaction as instruments for 3G coverage. More details about the IV approach can be found in Appendix D A.5.2.4.

5.5 Results

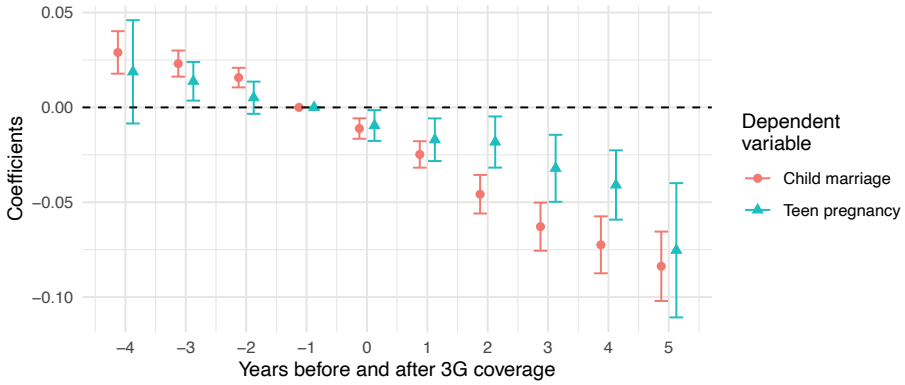
5.5.1 3G coverage, child marriage and teen pregnancy

Before presenting the results of the effect of 3G coverage on child marriage and teen pregnancy, we show the results of an event-study framework based on Sun and Abraham's (2021) approach that allows us to check for the lack of differential pre-trends, which is a precondition for the identification assumption of the TWFE and DiD models.

Figure 5.4 plots the point estimates of the event-study regression estimates for 3G coverage on child marriage and teen pregnancy. While adolescent girls with 3G had a slightly higher chance to get married in the years prior to the arrival of 3G with coefficients around 0.025, and a small decline in the coefficients is visible, the confidence intervals from years -4 until -2 overlap with each other but not with -1. Therefore, we cannot rule out a pre-trend for the child marriage analysis. Yet, we observe a stark decline starting from period 0. In response to the introduction of 3G, we observe a 15 decline in child marriage with some indication that it increases over time in absolute terms. Yet, confidence intervals for recent years are wide. Similarly, Figure 5.4 clearly shows the absence of a pre-trend for teen pregnancy. In the years prior to the roll-out of the 3G network, point estimates are statistically insignificant or close to zero with all confidence intervals overlapping. Starting with the

first year of 3G availability (t = 0), the point estimates become significantly negative, meaning that the availability of a 3G signal reduces the likelihood of young girls giving birth. Importantly, there is no fading of the effect as the magnitude of the effects increases in absolute terms in later years, implying a cumulative process.

Figure 5.4 Event study estimates of the impact of 3G on child marriage and teen pregnancy



Notes: The figure depicts the event study point estimates and the 95% confidence intervals of the dynamic effects of 3G coverage on child marriage and teen pregnancy for the Sun and Abraham (2021) estimator. For the child marriage estimates, we included all women U18 (N = 315,000); for the teen pregnancy estimates, we included all women U20 (N = 353,677). Results are produced with the full set of controls, i.e., woman-fixed effects and locational controls.

Table 5.2 shows the effect of 3G coverage on child marriage. We consistently find a significant negative effect of 3G coverage on child marriage. Column 1 shows that – with our most parsimonious TWFE model that only includes cluster and country × year fixed effects, the probability of child marriage decreases by 1.0 percentage point when an adolescent girl's cluster is covered by 3G. As we add additional control variables in Column 2 and 3, the effect remains stable. In our most demanding specification in Column 4, we add time-varying woman and location characteristics and women fixed effects and show that an adolescent girl's chances to become a child bride are 1.7 percentage points lower when she gains access to 3G coverage compared to no access. This is a considerable decline in the incidence of child marriage as it corresponds to an 8.2% decrease in child marriage evaluated at the sample average child marriage rate of 20.5%. If we consider the weighted stacked DID result (Column 5), the average treatment effect (ATT) stemming from Sun & Abraham's approach (Column 6) and the Borusyak estimation (Column 7), we obtain larger reductions of 2.9, 3.4, and 2.6 percentage points, respectively. However, this is the main effect of having 3 or 4 years of 3G coverage. The coefficients for the

first year are between 1.0 percentage points and 1.8 percentage points, and thus within the confidence intervals of the TWFE estimator (Appendix D A.5.2.1). The Local Average Treatment Effect estimated with the lightning IV is 21 percentage points, which translates into an 6.1 percentage points reduction in ITT terms, given the share of compliers of 28.8%. While larger than the effects estimated in the other models, this effect is still within the same order of magnitude.

The results for teen pregnancy are similar to those for child marriage and are presented in Table 5.3. Using a parsimonious version of our estimating equation (5.1) in which we only control for country x year and cluster fixed effects, the resulting regression coefficient presented in Column 1 shows that teen pregnancy reduces by 0.9 percentage points when a teenager is covered by 3G. The coefficient increases when we control for woman characteristics and 2G coverage to 1.0 percentage point in Column 2. In Column 3, we restrict the comparison to women who gained access to 3G by including woman-fixed effects. As such, we compare the probability of giving birth of the same adolescent girl, in the years before and after getting 3G coverage. The resulting coefficient of 1.8 percentage points is significant at the 1 percent level, and larger than the coefficients without woman-fixed effects. This coefficient slightly decreases to a 1.3 percentage points reduction in teen pregnancy when we include location controls in Column 4. This implies a 7.1% reduction in the probability of giving birth, relative to the control mean of 18.0%. Again, Columns 5 to 8 show that the estimates of the ATT of the weighted stacked DiD, Sun & Abraham's and Borusyak's model reflect 3 or 4 years of 3G coverage, and fall within the confidence intervals of the TWFE estimator. The LATE of the IV estimator shows a reduction in 5.23 percentage points for those covered by 3G, which translates into an ITT of 1.3 percentage points, which again falls within the confidence intervals of the TWFE estimator.

The TWFE results on both child marriage and teen pregnancy are robust to the inclusion of region x year trends and the inclusion of urban x year trends (Table A8), the exclusion of the open crowd information data of 3G and 4G (Table A9), and the exclusion of any country in a Jackknife approach (Figure A2).

Table 5.2: Effect of 3G coverage on child marriage

	Child marriage							
	TWFE model 1	**TWFE model 2**	**TWFE model 3**	**TWFE model 4**	**Weighted stacked DID**	**Sun & Abraham ATT**	**Borusyak et al. 2024**	**IV**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
M3G coverage	-0.0098***	-0.0091***	-0.0235***	-0.0168***	-0.0290***	-0.0335***	-0.0257***	-0.2114***
	(0.0014)	(0.0015)	(0.0020)	(0.0022)	(0.0027)	(0.0037)	(0.0020)	(0.0396)
Cluster FE	yes	yes	yes	yes	yes[1]	yes	no	no
Country year FE	yes	yes	yes	yes	yes[1]	yes	yes	no
Year FE	no	no	no	no	no	no	no	yes
Woman controls	no	yes	yes	yes	yes[1]	yes	yes	yes
Woman FE	no	no	yes	yes	yes[1]	yes	yes	yes
Locational controls	no	no	no	yes	yes[1]	yes	yes	yes
Obs.	490,946	435,924	435,924	352,352	352,187	918,036	277,474	315,000
R-squared	0.0977	0.1321	0.5299	0.5818	0.5468	0.6388		0.5862

Note: * p < .1; ** p < .05; *** p < .01; [1] fixed effects are cohort specific fixed effects. Standard errors (reported in parentheses) are clustered at the cluster level. Child marriage is 1 if a woman (under 18) got married in a given year and 0 otherwise. Woman controls include: 2G coverage and fixed effects of age. Locational controls include: frequency of negative and positive rainfall shocks, night-time lights, temperature, rainfall, and interactions between a linear time trend and elevation, distance to the coast, distance to a main river and distance to a city. Additional information can be found in appendix D A.5.2.1 for the weighted stacked DID estimation, A.5.2.2 for the Sun & Abraham method, A.5.2.3 for the Borusyak et al. 2024 estimation, and A.5.2.4 for the IV specification and first stage results.

Table 5.3: Effect of 3G coverage on teen pregnancy

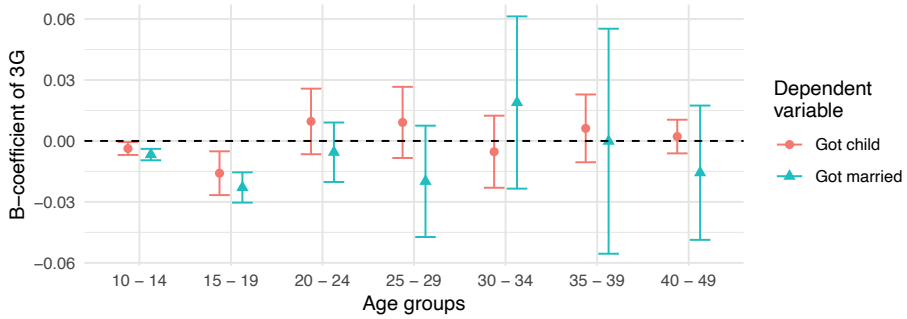
	Teen pregnancy							
	TWFE model 1	**TWFE model 2**	**TWFE model 3**	**TWFE model 4**	**Weighted stacked DID**	**Sun & Abraham ATT**	**Borusyak et al. 2024**	**IV**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
M3G coverage	-0.0087***	-0.0096***	-0.0182***	-0.0128***	-0.0195***	-0.0180***	-0.0188***	-0.0523*
	(0.0019)	(0.0019)	(0.0024)	(0.0031)	(0.0030)	(0.0049)	(0.0030)	(0.0295)
Cluster FE	yes	yes	yes	yes	yes[1]	yes	no	no
Country year FE	yes	yes	yes	yes	yes[1]	yes	yes	no
Year FE	no	no	no	no	no	no	no	yes
Woman controls	no	yes	yes	yes	yes[1]	yes	yes	yes
Woman FE	no	no	yes	yes	yes[1]	yes	yes	yes
Locational controls	no	no	no	yes	yes[1]	yes	yes	yes
Obs.	528,926	501,166	501,166	387,384	386,486	770,916	302,300	353,677
R-squared	0.0758	0.1312	0.4135	0.4476	0.4445	0.5652		0.4496

Note: * p < .1; ** p < .05; *** p < .01; [1] fixed effects are cohort specific fixed effects. Standard errors (reported in parentheses) are clustered at the cluster level. Teen pregnancy is 1 if a woman (under 20) got a child in a given year and 0 otherwise. Woman controls include: 2G coverage and fixed effects of age. Locational controls include: frequency of negative and positive rainfall shocks, night-time lights, temperature, rainfall, and interactions between a linear time trend and elevation, distance to the coast, distance to a main river and distance to a city. Additional information can be found in appendix D A.5.2.1 for the weighted stacked FE estimation, A.5.2.2 for the Sun & Abraham method, A.5.2.3 for the Borusyak et al. 2024 estimation, and A.5.2.4 for the IV specification and first stage results.

Next, we examine whether the negative effect of 3G on marriage and fertility is indeed specific to adolescent girls or whether there is an effect for older age groups as well. Figure 5.5 visualizes the regression coefficients from the TWFE for the effect of 3G coverage on the probability of getting married and giving birth for different age groups.

We find that the negative effect of 3G coverage on getting married and giving birth is strongest for women under the age of 20 and particularly for women aged 15-19. The effect of 3G coverage is also significantly negative for women aged 10-14 although the effect size is almost half of the effect size for women aged 15-19. We do not find an effect of 3G on the chances of getting married or giving birth for women aged 20 and older. This implies that it is certainly the youngest generation whose marriage and fertility patterns are affected by having access to 3G coverage, and that the younger generation is more susceptible to change their behavior following the introduction of 3G.

Figure 5.5 Effect of 3G coverage on marriage and fertility by age group



Notes: The Figure depicts the point estimates and 95% confidence intervals of the effect of 3G coverage on getting married and giving birth per age group using 14 different samples, 1 for each combination of age group and dependent variable. The results are derived with the TWFE model with the full set of controls, i.e., woman-fixed effects and locational controls that is also used for the estimates in Columns 4 in Table 2 and Table 3.

5.5.2 Heterogeneity analysis

Next, we test for heterogeneity in the effect of 3G coverage on child marriage and teen pregnancy by characteristics of the adolescent girl or her household. DiMaggio et al. (2001) and Norris (2001) identify several of these characteristics as important determinants of inequality in access to the internet. The results are presented in Figure 5.6. First, we test for heterogeneity by proxies for wealth of the household of the adolescent girl, namely by whether her household has access to electricity,

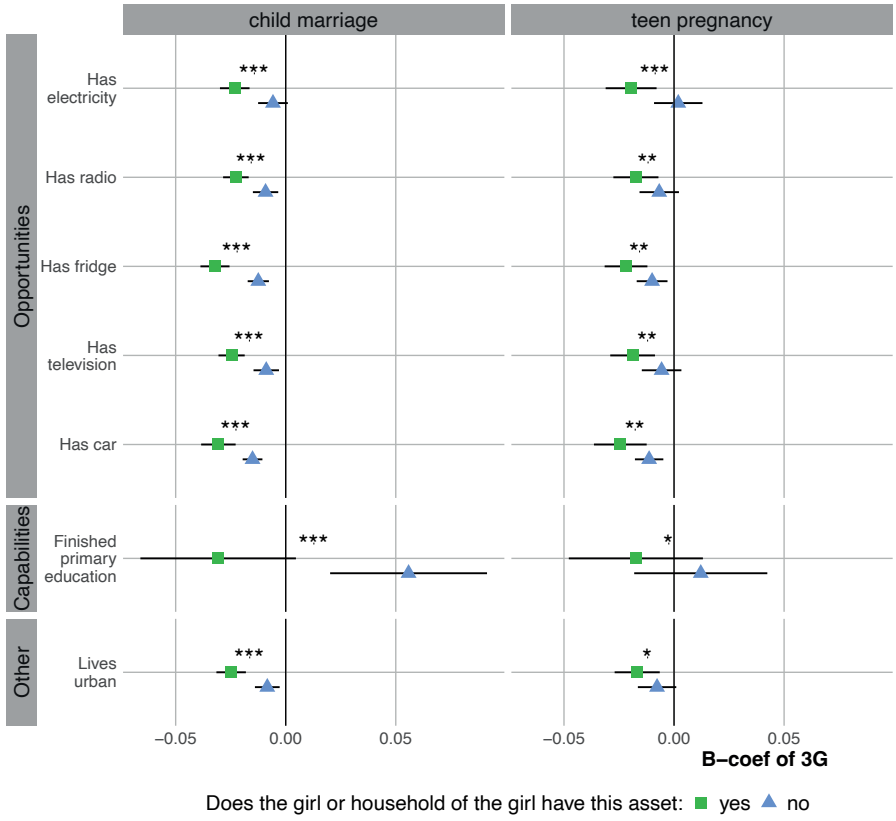
owns a radio, a fridge, a television, or a car. We also show heterogeneity by the adolescent girls' primary education, and we test for differences in effects between urban and rural settings.

All our models show that the negative effect of 3G is stronger for adolescent girls in households with electricity, a radio, a fridge, a television, and a car. This is the case for child marriage as well as teen pregnancy. We also find that the negative effect of 3G is almost completely driven by adolescent girls who finished their primary education. This suggests that the ability to read and write matters for the effect of 3G coverage, possibly because it promotes basic digital literacy skills, which are key for accessing and distilling information from mobile internet (Geldof et al., 2021). The effect of 3G on adolescent girls who have not finished primary education is opposite for child marriage (and increase in child marriage), and not significantly different from zero for teen pregnancy. We further find that the negative effect of 3G is more prominent in urban areas, which are the areas where women are more likely to have the opportunities and capabilities for using 3G. These results reinforce the earlier identified digital divide with respect to 3G technology.

5.5.3 Possible mechanisms

Our results so far document that the roll-out of 3G coverage reduces child marriage and teen pregnancy in SSA. We have, however, been silent about possible mechanisms underlying these reduced form effects. The 3G signal itself cannot directly cause child marriage and teen pregnancy to decline. Thus, access to 3G must be a trigger for changes in knowledge, attitudes or behavior that ultimately lead to the observed results. To shed some light on the differences in knowledge, attitudes and behavior of young women who have access and who don't have access to a 3G signal, we estimate a series of regression models with alternative dependent variables coming from the DHS and MICS surveys to examine whether increased access to information and changes in attitudes and preferences may be the drivers of the effect on child marriage and teen pregnancy we identify. The data, along with the regression specifications are discussed in more detail in Appendix D A.5.4. All subsequently presented results should be interpreted as suggestive, as the available data do not allow for the same rigor in the analysis as in the child marriage and teen pregnancy specifications since we only have data available for one moment in time, i.e., the moment of the survey.

Figure 5.6 Effect of 3G coverage by adolescent girls' characteristics



Notes: * p < .1; ** p < .05; *** p < .01. The Figure depicts the regression estimates and the 95% confidence interval of the effect of 3G coverage for having (= green yes) and not having (= blue no) certain assets. The upper line stars show whether these differences are significant. Results are derived with the full set of controls, i.e., woman-fixed effects and locational controls. Child marriage is 1 if a woman under 18 got married in a given year and teen pregnancy is 1 if a woman under 20 gave birth in a given year and 0 otherwise. Table A10 shows the complete regression results for child marriage, Table A11 shows them for teen pregnancy.

We start by examining whether 3G had an effect on information. We assess whether 3G coverage is a predictor of adolescent girls' knowledge about modern contraceptives, awareness of family planning programs, and AIDS and sexually transmittable infections (STIs). Table 5.4 shows that, consistent with the idea that mobile broadband internet facilitates the exchange of information, we find evidence for differences in knowledge. Young women are better informed about modern contraceptives, family planning programs, and AIDS and STIs when they are covered by 3G.

Next, we address whether adolescent girls change their attitudes and preferences concerning gender norms, preferred marriage behavior, and fertility in response to 3G coverage. In particular, we test whether 3G coverage predicts adolescent girls' acceptance towards wife beating and refusing sex, and their ideal family size. As shown in Table 5.5, we find that 3G coverage makes women more likely to report that they find wife beating unacceptable, they are more likely to report that refusing sex is acceptable, and they report a lower ideal family size.

Finally, we consider effects of 3G coverage on behavior. Column 1 in Table 5.6 shows that we do not find an effect of 3G coverage on the onset of contraceptive use (the age at which they use contraceptives for the first time). Column 2 shows, however, that 3G coverage has a significant positive effect on whether or not adolescent girls are currently using contraceptives. We also find that there is a decreased likelihood that women report to have started having sexual intercourse when they have 3G coverage (Column 3). Although this is only one facet of women's sexual and family formation, it is one effective way of avoiding teen pregnancy. It is also related to the finding that contraceptive use does not start earlier since girls who do not engage in sexual relationships are unlikely to use contraceptives. Thus, while young women are not starting the use of contraceptives at a younger age, they are more likely to keep using them when covered by 3G. This is an important finding as family planning programs often struggle with expanding contraceptive use.

Table 5.4: Effect of 3G on knowledge

	Knowledge channel		
	Knowledge of modern contraceptives	**Heard about family planning**	**Heard of Aids or other STI**
M3G coverage	0.0125***	0.0346***	0.0193***
	(0.0047)	(0.0113)	(0.0073)
Region FE	yes	yes	yes
Country year FE	yes	yes	yes
Woman controls	yes	yes	yes
Locational controls	yes	yes	yes
Obs.	42,712	39,085	38,357
R-squared	0.1346	0.2289	0.4033
Mean of dep. var.	0.9231	0.4719	0.778

Note: * p < .1; ** p < .05; *** p < .01. All women under 20 are included on which we have information on the dependent variable using only the most recent observation per woman. Standard errors (reported in parentheses) are clustered at the region level. Woman controls are the age of the respondent, whether the household has a car, television, fridge, electricity and whether the woman finished primary education. Locational controls include minimum and maximum temperature, rainfall, nightlights, elevation, ocean distance, river distance, travel times to the nearest city, and negative and positive rain shocks. More details on the estimation strategy and variables are provided in appendix D A.5.4.

Table 5.5: Effect of 3G on attitudes and preferences

	Attitudes and preference channel		
	Wife beating unacceptable	**Refusing sex accepted**	**Ideal family size**
M3G coverage	0.0752***	0.0445**	-0.2217***
	(0.0201)	(0.0190)	(0.0358)
Region FE	yes	yes	yes
Country year FE	yes	yes	yes
Woman controls	yes	yes	yes
Locational controls	yes	yes	yes
Obs.	64,668	7,990	37,667
R-squared	0.1352	0.2198	0.2845
Mean of dep. var.	3.2683	0.5894	4.2064

Note: * p < .1; ** p < .05; *** p < .01. All women under 20 are included on which we have information on the dependent variable using only the most recent observation per woman. Standard errors (reported in parentheses) are clustered at the region level. Woman controls are the age of the respondent, whether the household has a car, television, fridge, electricity and whether the woman finished primary education. Locational controls include minimum and maximum temperature, rainfall, nightlights, elevation, ocean distance, river distance, travel times to the nearest city, and negative and positive rain shocks. More details on the estimation strategy and variables are provided in appendix D A.5.4.

Table 5.6: Effect of 3G on behavior

	Behavioral effects		
	Onset contraceptive usage	**Current contraceptive usage**	**Onset sexual intercourse**
M3G coverage	-0.0005	0.0088*	-0.0131***
	(0.0125)	(0.0050)	(0.0036)
Cluster FE	yes	no	yes
Region FE	no	yes	no
Country year FE	yes	yes	yes
Woman controls	yes	yes[1]	yes
Woman FE	yes	no	yes
Locational controls	yes	yes	yes
Obs.	56,128	42,712	449,161
R-squared	0.5476	0.1177	0.8849
Mean of dep. var.	0.0538	0.2588	0.4983

Note: * p < .1; ** p < .05; *** p < .01. All women under 20 are included on which we have information on the dependent variable. Onset contraceptive usage and sexual intercourse are panel datasets as used for our child marriage models only now with a different dependent variable, current contraceptive usage is a cross-sectional dataset including only the most recent observation of the woman. Standard errors (reported in parentheses) are clustered at the cluster level for onset contraception usage and sexual intercourse and at the regional level for current contraceptive usage. Onset contraceptive usage is 1 if a woman reports the start of using it and 0 otherwise. Current contraceptive usage is 1 if

a woman is currently using. Sexual intercourse is 1 if a woman had for the first time sexual intercourse. Woman controls include: 2G coverage and fixed effects of age, and woman controls[1] are the age of the respondent, whether the household has a car, television, fridge, electricity and whether the woman finished primary education. Locational controls include: frequency of negative and positive rainfall shocks, night-time lights, temperature, rainfall, and interactions between a linear time trend and elevation, distance to the coast, distance to a main river and distance to a city.

5.6 Discussion and conclusion

We established a robust negative relationship between access to mobile broadband internet and child marriage and teen pregnancy in SSA. The effects of 3G are particularly pronounced for young women who have access to electricity, those who have higher wealth, live in urban areas and who have finished primary education.

The economic magnitudes of the effects of 3G on child marriage and teen pregnancy that we find are substantial: an 8% decrease in the likelihood of child marriage and a 7% decrease in the likelihood of teen pregnancy. Comparing our effect sizes with earlier studies that looked at the effect of ICTs and media, we see that our results align. Ferrara et al. (2012) reported a 7% decline in fertility due to the Brazilian telenovelas, Guldi and Herbst (2017) identified a 7% decline in teen fertility due to broadband internet in the United States, Kearney and Levine (2015) a 4% reduction in fertility due to MTV's TV show 16 and pregnant, and lastly, Bönisch and Hyll (2023) a 4-6% reduction in fertility due to the arrival of television.

We show that potential drivers of the effect of 3G coverage on child marriage and teen pregnancy are knowledge of contraceptives, STIs, and family planning; less favorable attitudes towards repressive behavior against women, a lower desired fertility; and a decrease in the likelihood that women report that they have started sexual intercourse, and a higher likelihood of current contraceptive use. While we focus on the knowledge-attitude-behavior nexus because we have the most credible data to test this mechanism, we do not attempt to give a complete picture of all possible channels through which 3G might influence child marriage and teen pregnancy. Other studies found that having access to 2G increased women's labor market participation and income, which has the potential to protect women against harmful behaviors such as child marriage and teen pregnancy (Bahia et al., 2024); 3G is likely to have the same effect. It is known that having access to mobile money (through the 2G network) serves as protection against income shocks and as such, may serve as a protection mechanism for child marriage in certain contexts (Corno et al., 2020; Corno & Voena, 2023).

Our study is not without limitations. First, we rely on observational data. Yet, we can draw on a large sample with a considerable geographical and temporal spread. Moreover, the underlying surveys are representative of the population. Second, the staggered introduction of 3G poses the challenge of making forbidden comparisons. This is why we complement our TWFE with the weighted stacked DiD. For the child marriage analyses we cannot exclude the presence of pre-trends, therefore we exploit conditional random variation in lighting as an instrument for 3G coverage, to estimate the effect on teen fertility and child marriage. Also, the IV model that uses the frequency of lightning strikes has challenges as lightning is time invariant and needs to be interacted with time. While the relevance of this instrument has been established by previous studies, we observe that the relationship between the lightning-trend and 3G coverage is not linear in our data. Therefore, we use the lightning-trend and its squared term, to allow its effect to be more flexible. Third, our study of underlying mechanisms is only exploratory due to the lack of higher quality data. More research is needed on underlying drivers of the effects we observe.

We want to conclude with a word of caution. Our findings should not be interpreted in an overly optimistic spirit, inferring that access to mobile broadband internet will necessarily empower all young women in SSA. Our heterogeneity analysis shows that benefits from 3G coverage are heterogeneous. As access to mobile broadband internet grows, the disadvantage of not having a compatible device for different social and economic reasons may increase existing inequalities. Girls who do not have access to a phone might get one from a man/boy, but such gifts often include return favors rooted in traditions of patriarchy (Porter et al., 2020). On the other hand, externalities from girls with access to mobile broadband, through information, norms or (marriage) market dynamics could improve outcomes for girls without access to broadband. Moreover, girls searching for information on, for example, family planning, may encounter misinformation, suggesting that modern contraceptives cause infertility at older age, a claim widely disseminated in Africa (Boivin et al., 2020; Engelbert Bain et al., 2021). Similarly, social media does not automatically lead to improved decision-making, as individuals can self-select content based on their prior preferences and attitudes, leading to echo chambers.

The lesson learned from this paper is that access to mobile broadband internet has been serving as an empowerment tool for young women. We find an economically substantial effect, especially if we compare our results with targeted family planning programs that often struggle to reach substantial results (Feyissa et al., 2023). While program designers of family planning interventions who aim to

empower adolescent girls are advised to address environmental factors, including financial barriers and community norms, 3G coverage seems to indirectly affect and alter these factors. As such, investing in 3G coverage, together with investments in socioeconomically structured opportunities and capabilities for girls who are at risk of falling behind, seems a promising policy strategy to reduce child marriage and teen pregnancy, and ultimately empower adolescent girls in SSA. Thus, programs that target women empowerment and family planning would benefit from incorporating the use of mobile phones and internet in their programs.

In sum, this paper shows that the expansion of mobile broadband internet substantially reduced child marriage and teen pregnancy in SSA. The findings imply that a business-led endeavor, the expansion of mobile broadband coverage, can be a stepping stone for women's empowerment in developing countries. This calls for more creative approaches to development collaboration and for strategic partnerships between governments, donors, and businesses on the route to achieving SDG5 of gender equality by 2030.

Chapter 6: Conclusion

6.1 Research questions revisited

This dissertation set out to answer four overarching research questions:

1. To what extent and in what ways do fertility preferences in SSA vary between ethnic groups? And how can this variation be explained by socioeconomic, demographic, and cultural characteristics of ethnic groups?
2. How does the strength of social norms, as captured by variation in the timing of marriage and first birth within social groups, affect young women's incidence of early marriage and adolescent childbearing in SSA?
3. To what extent is regional variation in crude birth rates associated with levels of social media usage and gender gaps in digital access across SSA?
4. How does the expansion of mobile broadband (3G) coverage influence the incidence of child marriage and teen pregnancy in SSA?

On the first question, the analysis of more than 500,000 women across 181 ethnic groups in 24 countries showed that fertility preferences vary substantially between ethnic groups. These differences remained even after controlling for women's education, household wealth, and other individual factors, underscoring the importance of ethnicity as a cultural framework. Group-level characteristics such as average education and gender-related norms further explained variation, and their effects were moderated by women's own resources, such as education and occupation.

Regarding the second question, the results show that the strength of social norms, proxied by within-group variation in the timing of family formation, has a strong influence on early transitions. In groups with little variation (strong norms), marriage and childbirth occur earlier, while in groups with greater variation (weaker norms), women marry and give birth later. A one standard deviation increase in variation is associated with a delay of 3–4 months in marriage and 4–5 months in first birth, and a nearly two-percentage-point reduction in the likelihood of child marriage and teen pregnancy. Norms transmitted by older generations appear to matter more for marriage timing, while fertility is more strongly influenced by peers.

For the third question, analyses of 311 subnational regions in 29 countries reveal a robust negative relationship between social media usage and fertility outcomes. Higher Facebook penetration and smaller gender gaps in access are associated with significantly lower crude birth rates, even when accounting for education, wealth, and urbanization. These findings suggest that social media facilitates the diffusion

of alternative reproductive norms, while persistent gender gaps in access sustain higher fertility.

On the fourth question, exploiting the staggered roll-out of 3G coverage between 2011 and 2021 across 20 countries shows that mobile internet expansion has measurable effects on early family formation. Exposure to 3G reduced the probability of child marriage by 1–6 percentage points and of teen pregnancy by around 1 percentage point. These effects are most pronounced among wealthier and better-educated girls, indicating that benefits are unevenly distributed. Mechanism analyses suggest that increased access to information, changes in gender attitudes, and greater contraceptive use are key channels through which digital connectivity influences outcomes.

6.2 Conclusion

The findings revealed the importance of social structures and the role of digital infrastructure in shaping women's timing of family formation and family size preferences in SSA. Across diverse contexts, the findings demonstrate that reproductive decisions are not only the result of individual preferences or resource constraints but are also deeply embedded in the social environments in which women live. This aligns with a long tradition of demographic research showing that fertility behavior is strongly shaped by cultural and social contexts (Bongaarts & Watkins, 1996; Caldwell & Caldwell, 1987; Johnson-Hanks, 2005; Trinitapoli & Yeatman, 2018). Norms linked to ethnicity, religion, and local community structures define what is considered appropriate in terms of marriage timing, childbearing, and desired family size (Ahonsi et al., 2019; Mkhwanazi, 2010; Schaffnit et al., 2019). As also emphasized by Casterline (2017) and Batyra (2021), such norms can sustain high fertility and early transitions even when education, economic conditions, and legal protections improve.

At the same time, the rapid expansion of digital connectivity - particularly mobile internet and social media - has introduced new informational landscapes that can challenge traditional norms (Ferrara et al., 2012; Manacorda & Tesei, 2020; Rotondi et al., 2020). Digital platforms expose women and girls to alternative role models, diverse life trajectories, and information about reproductive health and rights, which can expand the perceived range of acceptable choices (Fatehkia et al., 2018; Kearney & Levine, 2015). This thesis shows that such exposure is linked to later marriage, lower fertility, and a reduction in early childbearing, although these

benefits are not equally distributed. Digital opportunities tend to be taken up more readily by women who are already relatively advantaged, such as those with higher education or living in urban areas, highlighting persistent inequalities in access and impact.

Together, the results depict a layered reality: women's reproductive behaviors are shaped by the interplay of strong, place-based social norms and increasingly influential digital networks. Local normative structures continue to exert powerful influence, but digital infrastructure offers channels for normative pluralism, where competing expectations and new aspirations can coexist. Understanding this interaction is crucial for designing policies that both acknowledge the weight of social context and leverage the transformative potential of digital connectivity to promote reproductive autonomy and gender equality.

6.3 Policy implications

The findings of this thesis point to several priorities for policy aimed at delaying early marriage and childbearing, reducing high fertility, and promoting women's reproductive autonomy in SSA. Interventions that focus solely on individual-level empowerment - such as scholarships for girls or awareness campaigns - risk limited impact if they do not address the normative environments in which decisions are made. Programs should actively involve local leaders, elders, religious authorities, and peer influencers in shaping new narratives about appropriate marriage timing and family size. Building trust and fostering community dialogue can help shift social expectations in ways that make delayed marriage and smaller families socially acceptable.

The evidence that low variation in marriage and childbirth timing predicts earlier transitions suggests that "tight" normative environments require different strategies from more flexible ones. In high-pressure settings, policies may need to prioritize collective behavior change - through community pledges, public commitments, or alternative rites of passage - rather than focusing only on individual choices. In low-pressure settings, policies can be more effectively targeted toward individual women, as they are likely to have greater autonomy in their family formation choices.

Mobile internet and social media can be powerful tools for expanding young women's access to reproductive health information, alternative role models, and

supportive networks. Partnerships between governments, NGOs, and telecom providers could expand affordable access, especially in rural areas, while ensuring that digital content is accurate, relevant, and culturally sensitive. Digital literacy initiatives can help young women navigate online spaces safely and critically. The positive associations between digital access and delayed family formation are strongest among more educated and wealthier women. Without deliberate action, digital expansion risks widening existing inequalities, making it essential to prioritize affordable access, device provision, and skills training for girls and women in low-income, rural, and marginalized communities.

Digital interventions are most effective when reinforced by offline support systems. Linking online information campaigns to in-person services - such as youth-friendly health clinics, counselling, and safe spaces - can help translate new knowledge and aspirations into actual behavior change. Given the dynamic interaction between local norms and digital exposure, policies should invest in monitoring systems that track both behavioral and normative indicators over time. This can inform adaptive programming that responds to shifts in attitudes, community discourse, and digital engagement patterns.

By simultaneously addressing the persistence of local social norms and the opportunities opened by digital connectivity, policymakers can design interventions that are both context-sensitive and forward-looking. Such strategies hold the potential not only to reduce early marriage and high fertility, but also to expand the life choices available to women and girls across the region.

6.4 Limitations

Several limitations should be acknowledged when interpreting the findings of this thesis. First, while the analyses demonstrate strong associations between social group characteristics and women's reproductive outcomes, they cannot fully capture how deeply individuals feel embedded within these social structures. The data do not provide measures of subjective identification with religion, ethnicity, or local community, nor do they indicate the relative weight of these affiliations in shaping behavior. As a result, the extent to which observed group-level effects reflect active normative influence versus passive social background remains uncertain.

Second, the measures of digital connectivity capture access and broad usage patterns rather than the nature and intensity of engagement. For example, Facebook usage rates and mobile broadband coverage indicate potential exposure to online information, but they do not reveal what content people consume, how much time they spend online, or whether interactions challenge or reinforce existing norms. Without detailed usage data which can be linked to individual survey data, it is not possible to disentangle the effects of mere connectivity from those of specific online behaviors or networks.

Third, limitations in the available survey and geospatial data constrain the precision of measurement. Smaller or marginalized groups are sometimes excluded due to insufficient sample size, potentially biasing results toward larger or more visible communities. Digital trace data may overrepresent younger, urban, and more educated populations, while coverage maps may overstate actual access if infrastructure is present but devices, affordability, or digital literacy remain barriers.

Fourth, while quasi-experimental methods in some analyses strengthen causal inference, unobserved confounding cannot be ruled out. Broader developmental changes, infrastructure investments, or policy shifts occurring alongside the spread of digital technologies may have influenced the observed outcomes.

Finally, the study infers mechanisms - such as increased access to reproductive health information, exposure to alternative role models, and changes in gender attitudes - from patterns in the data rather than observing them directly. Without direct evidence of these processes, conclusions about how digital access translates into behavioral change remain suggestive, and future work should seek to measure these pathways more explicitly.

6.5 Future research

Building on the findings and limitations of this thesis, future research should aim to deepen the understanding of how social structures and digital connectivity interact to shape women's family formation in Sub-Saharan Africa. A priority is to capture how strongly individuals identify with different social groups—such as ethnic, religious, or neighborhood communities—and how this identification shapes their responsiveness to group norms. Surveys incorporating measures of subjective belonging, perceived normative pressure, and the salience of various identities

would make it possible to distinguish between passive group membership and active normative influence.

Further work should also investigate online engagement in more detail. Beyond measuring access, it is important to know what people do online, what content they encounter, who they interact with, and how much time they spend in these digital environments. Linking digital trace data with survey responses or time-use diaries could help identify the pathways through which online exposure translates into changes in attitudes, aspirations, and behaviors.

To address the temporal dynamics of change, longitudinal and panel data would be invaluable. These designs could capture the sequencing of digital adoption, normative shifts, and reproductive decisions, allowing researchers to observe both immediate and lagged effects. Such data could also be used to track the longer-term consequences of delayed marriage or reduced adolescent childbearing for women's education, employment, and well-being.

Finally, mixed-methods approaches that combine large-scale quantitative analyses with qualitative insights from interviews, focus groups, and ethnographic work could reveal how women negotiate competing offline and online influences in their everyday lives. This would enrich understanding of the mechanisms behind observed statistical patterns and help identify which types of digital and community-based interventions are most effective in specific contexts.

Through these extensions, future research can move toward a more complete account of how entrenched local norms and expanding digital networks jointly influence reproductive autonomy and demographic change in the region.

References

Abubakar, N. H., & Dasuki, S. I. (2018). Empowerment in their hands: Use of WhatsApp by women in Nigeria. *Https://Doi.Org/10.1080/09718524.2018.1509490*, *22*(2), 164–183. https://doi.org/10.1080/09718524.2018.1509490

Abu-Shanab, E., & Al-Jamal, N. (2015). Exploring the Gender Digital Divide in Jordan. *Gender, Technology and Development*, *19*(1), 91–113. https://doi.org/10.1177/0971852414563201

Addai, I. (1999). ETHNICITY AND CONTRACEPTIVE USE IN SUB-SAHARAN AFRICA: THE CASE OF GHANA. *Journal of Biosocial Science*, *31*(1), 105–120. https://doi.org/10.1017/S0021932099001054

Adedini, S. A., Odimegwu, C., Imasiku, E. N. S., & Ononokpono, D. N. (2015). Ethnic differentials in under-five mortality in Nigeria. *Ethnicity & Health*, *20*(2), 145–162. https://doi.org/10.1080/13557858.2014.890599

Afulani, P. A., & Asunka, J. (2015). Socialization, Adaptation, Transnationalism, and the Reproductive Behavior of Sub-Saharan African Migrants in France. *Population Research and Policy Review*, *34*(4), 561–592. https://doi.org/10.1007/s11113-015-9360-2

Agege, E. A., Nwose, E. U., & Odjimogho, S. (2018). Parental perception of girl-child early marriage amongst the Urhobos in Nigeria. *International Journal of Community Medicine and Public Health*, *5*(8), 3185–3190.

Ahonsi, B., Fuseini, K., Nai, D., Goldson, E., Owusu, S., Ndifuna, I., Humes, I., & Tapsoba, P. L. (2019). Child marriage in Ghana: Evidence from a multi-method study. *BMC Women's Health*, *19*, 1–15.

Airhihenbuwa, C. O. (2007). *Healing our differences: The crisis of global health and the politics of identity*. Bloomsbury Publishing PLC.

Aker, J. C., & Mbiti, I. M. (2010). Mobile phones and economic development in Africa. *Journal of Economic Perspectives*, *24*(3), 207–232.

Akerlof, G. A. (1980). A Theory of Social Custom, of Which Unemployment May be One Consequence. *The Quarterly Journal of Economics*, *94*(4), 749–775.

Akerlof, G. A., & Kranton, R. E. (2000). Economics and Identity*. *The Quarterly Journal of Economics*, *115*(3), 715–753. https://doi.org/10.1162/003355300554881

Alaba, O. O., Olubusoye, O. E., & Olaomi, J. (2017). Spatial patterns and determinants of fertility levels among women of childbearing age in Nigeria. *South African Family Practice*, *59*(4), 143–147. https://doi.org/10.1080/20786190.2017.1292693

Alexander, M., Polimis, K., & Zagheni, E. (2019). The Impact of Hurricane Maria on Out-migration from Puerto Rico: Evidence from Facebook Data. *Population and Development Review*, *45*(3), 617–630. https://doi.org/10.1111/padr.12289

Allison, P. D. (2009). *Fixed Effects Regression Models*. SAGE Publications, Inc.

Alozie, N. O., Akpan-Obong, P., Fatehkia, M., Kashyap, R., & Weber, I. (2018). Using Facebook ad data to track the global digital gender gap. *Development Policy Review*, *107*(2), 189–209.

Arias, E. (2019). How does media influence social norms? Experimental evidence on the role of common knowledge. *Political Science Research and Methods*, *7*(3), 561–578.

Asadullah, M. N., & Wahhaj, Z. (2019). Early Marriage, Social Networks and the Transmission of Norms. *Economica*, *86*(344), 801–831. https://doi.org/10.1111/ecca.12291

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*(1), 54. https://doi.org/10.1186/s12905-019-0747-9

Austin, P. C., Stryhn, H., Leckie, G., & Merlo, J. (2018). Measures of clustering and heterogeneity in multilevel Poisson regression analyses of rates/count data. *Statistics in Medicine*, *37*(4), 572–589. https://doi.org/10.1002/sim.7532

Bahia, K., Castells, P., Cruz, G., Masaki, T., Pedrós, X., Pfutze, T., Rodríguez-Castelán, C., & Winkler, H. (2024). The welfare effects of mobile broadband internet: Evidence from Nigeria. *Journal of Development Economics*, *170*, 103314.

Bakibinga, P., Mutombo, N., Mukiira, C., Kamande, E., Ezeh, A., & Muga, R. (2016). The Influence of Religion and Ethnicity on Family Planning Approval: A Case for Women in Rural Western Kenya. *Journal of Religion and Health*, *55*(1), 192–205. https://doi.org/10.1007/s10943-015-0030-9

Balbo, N., & Barban, N. (2014). Does fertility behavior spread among friends? *American Sociological Review*, *79*(3), 412–431.

Banerjee, A., La Ferrara, E., & Orozco, V. (2019). Entertainment, education, and attitudes toward domestic violence. *American Economic Review, Papers and Proceedings*, *109*, 133–137.

Bates, R. H. (2000). Ethnicity and Development in Africa: A Reappraisal. *American Economic Review*, *90*(2), 131–134. https://doi.org/10.1257/aer.90.2.131

Batyra, E., Kohler, H.-P., & Furstenberg, F. F. (2021). Changing gender gaps in the timing of first union formation and sexual initiation in sub-Saharan Africa. *Population and Development Review*, *47*(2), 289–322.

Bau, N., Henning, D. J., Low, C., & Steinberg, B. (2024). *Family Planning, Now and Later: Infertility Fears and Contraceptive Take-Up* (Working Paper No. 32735; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w32735

Bell, A., Fairbrother, M., & Jones, K. (2019). Fixed and random effects models: Making an informed choice. *Quality & Quantity*, *53*(2), 1051–1074.

Bellemare, M., Novak, L., & Steinmetz, T. L. (2015). All in the family: Explaining the persistence of female genital cutting in West Africa. *Journal of Development Economics*, *116*(C), 252–265.

Bernardi, L. (2003). Channels of Social Influence on Reproduction. *Population Research and Policy Review*, *22*(5), 427–555. https://doi.org/10.1023/B:POPU.0000020892.15221.44

Bernardi, L., Keim, S., & von der Lippe, H. (2007). Social Influences on Fertility: A Comparative Mixed Methods Study in Eastern and Western Germany. *Journal of Mixed Methods Research*, *1*(1), 23–47. https://doi.org/10.1177/2345678906292238

Bernardi, L., & Klärner, A. (2014). Social networks and fertility. *Demographic Research*, *30*, 641–670. https://doi.org/10.4054/DemRes.2014.30.22

Berndt, J. O., Rodermund, S. C., & Timm, I. J. (2018). SOCIAL CONTAGION OF FERTILITY: AN AGENT-BASED SIMULATION STUDY. *2018 Winter Simulation Conference (WSC)*, 953–964. https://doi.org/10.1109/WSC.2018.8632331

Bertocchi, G., Dimico, A., & Falco, C. (2024). *Family Planning and Ethnic Heritage: Evidence from Sub-Saharan Africa*. IZA Discussion Papers.

Bicchieri, C. (2006). *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press. https://doi.org/10.1017/CBO9780511616037

Bicchieri, C., Jiang, T., & Lindemans, J. W. (2014). A social norms perspective on child marriage: The general framework. *Working Paper Penn Social Norms Group (PennSoNG)*.

Billari, F. C., Rotondi, V., & Trinitapoli, J. (2020). Mobile phones, digital inequality, and fertility. *Demographic Research*, *42*, 1057–1096.

Boivin, J., Carrier, J., Zulu, J. M., & Edwards, D. (2020). A rapid scoping review of fear of infertility in Africa. *Reproductive Health*, *17*, 1–13.

Bongaarts, J. (2009). Human population growth and the demographic transition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1532), 2985–2990.

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. (2017). Africa's Unique Fertility Transition. *Population and Development Review*, *43*(S1), 39–58. https://doi.org/10.1111/j.1728-4457.2016.00164.x

Bongaarts, J. (2020). Trends in fertility and fertility preferences in sub-Saharan Africa: The roles of education and family planning programs. *Genus*, *76*(1), 32. https://doi.org/10.1186/s41118-020-00098-z

Bongaarts, J., & Casterline, J. (2013). Fertility Transition: Is sub-Saharan Africa Different? *Population and Development Review*, *38*, 153–168.

Bongaarts, J., & Watkins, S. C. (1996). Social interactions and contemporary fertility transitions. *Population and Development Review*, 639–682.

Bönisch, P., & Hyll, W. (2023). Television and fertility: Evidence from a natural experiment. *Empirical Economics*, *64*(3), 1025–1066.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, rdae007.

Bourdieu, P. (1986). The Forms of Capital. In J. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241–258). Greenwood.

Branson, N., & Byker, T. (2018). Causes and consequences of teen childbearing: Evidence from a reproductive health intervention in South Africa. *Journal of Health Economics*, *57*, 221–235.

Burgert, C. R., Colston, J., Roy, T., & Zachary, B. (2013). *Geographic displacement procedure and georeferenced data release policy for the Demographic and Health Surveys*. Icf International.

Bursztyn, L., & Jensen, R. (2017). Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure. *Annual Review of Economics*, *9*(1), 131–153.

Caldwell, J. C., & Caldwell, P. (1987). The Cultural Context of High Fertility in sub-Saharan Africa. *Population and Development Review*, *13*(3), 409–437. JSTOR. https://doi.org/10.2307/1973133

Caldwell, J. C., & Caldwell, P. (2002). Africa: The new family planning frontier. *Studies in Family Planning*, *33*(1), 76–86.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.

Cassidy, R., Dam, A., Janssens, W., Kiani, U., & Morsink, K. (2024). *Targeting men, women or both to reduce child marriage* (Working Paper No. 24/23). Institue for Fiscal studies. https://doi.org/10.1920/wp.ifs.2024.2324

Casterline, J. B. (2017). Prospects for Fertility Decline in Africa. *Population and Development Review*, *43*, 3–18.

Castro Torres, A. F., Pesando, L. M., Kohler, H.-P., & Furstenberg, F. (2022). Family change and variation through the lens of family configurations in low-and middle-income countries. *Population, Space and Place*, *28*(4), e2531.

Cesare, N., Lee, H., McCormick, T., Spiro, E., & Zagheni, E. (2018). Promises and Pitfalls of Using Digital Traces for Demographic Research. *Demography*, *55*(5), 1979–1999. https://doi.org/10.1007/s13524-018-0715-2

Channon, M. D., & Harper, S. (2019). Educational differentials in the realisation of fertility intentions: Is sub-Saharan Africa different? *PLOS ONE*, *14*(7), e0219736. https://doi.org/10.1371/journal.pone.0219736

Chari, A., Heath, R., Maertens, A., & Fatima, F. (2017). The causal effect of maternal age at marriage on child wellbeing: Evidence from India. *Journal of Development Economics*, *127*, 42–55.

Child Marriage Data Portal. (2024). *Are we on track to end child marriage by 2030?*

Chiplunkar, G., & Goldberg, P. K. (2022). *The Employment Effects of Mobile Internet in Developing Countries* (Working Paper No. 30741; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w30741

Cleland, J., Bernstein, S., Ezeh, A., Faundes, A., Glasier, A., & Innis, J. (2006). Family planning: The unfinished agenda. *Lancet (London, England)*, *368*(9549), 1810–1827. https://doi.org/10.1016/S0140-6736(06)69480-4

Coale, A. J., & Watkins, S. C. (1986). *The Decline of Fertility in Europe*. 484.

Cole, W. M., & Geist, C. (2021). Conceiving of contraception: World society, cultural resistance, and contraceptive use, 1970–2012. *Social Forces*, *99*(4), 1394–1431.

Corno, L., Hildebrandt, N., & Voena, A. (2020). Age of marriage, weather shocks, and the direction of marriage payments. *Econometrica*, *88*(3), 879–915.

Corno, L., & Voena, A. (2016). *Selling daughters: Age of marriage, income shocks and the bride price tradition*. IFS Working Papers.

Corno, L., & Voena, A. (2023). Child marriage as informal insurance: Empirical evidence and policy simulations. *Journal of Development Economics*, 103047.

Coyne, C., & Coyne, R. L. (2014). The identity economics of female genital mutilation. *Journal of Developing Areas*, *48*(2), 137–152.

Dewi, R. K., Suryadarma, D., & Suryahadi, A. (2018). The impact of media on behaviour: Evidence from television coverage expansion and declining fertility in Indonesia. *Development Policy Review*, *36*, O552–O563.

Dibaba, B., & Mitike, G. (2016). Factors influencing desired family size among residents of Assela Town. *J Women's Health Care*, *5*(6), 1–8.

DiMaggio, P., Hargittai, E., Neuman, W. R., & Robinson, J. P. (2001). Social implications of the Internet. *Annual Review of Sociology*, *27*(1), 307–336.

DiMaggio, P., Hargittai, E., Russell Neuman, W., & Robinson, J. P. (2001). Social implications of the internet. *Annual Review of Sociology*, *27*, 307–336. https://doi.org/10.1146/annurev.soc.27.1.307

Easterlin, R. A., & Crimmins, E. M. (1985). *The fertility revolution: A supply-demand analysis*. University of Chicago Press.

Easterly, W., & Levine, R. (1997). Africa's Growth Tragedy: Policies and Ethnic Divisions*. *The Quarterly Journal of Economics*, *112*(4), 1203–1250. https://doi.org/10.1162/003355300555466

Eker, S., Garcia, D., Valin, H., & Van Ruijven, B. (2021). Using social media audience data to analyse the drivers of low-carbon diets. *Environmental Research Letters*, *16*(7), 074001.

Elster, J. (1989). *The Cement of Society: A Study of Social Order*. Cambridge University Press. https://www.cambridge.org/core/books/cement-of-society/2F5C6A09E882A9D54A4E4973A792CFE1

Engelbert Bain, L., Amu, H., & Enowbeyang Tarkang, E. (2021). Barriers and motivators of contraceptive use among young people in Sub-Saharan Africa: A systematic review of qualitative studies. *PloS One*, *16*(6), e0252745.

Ezeh, A. C., Mberu, B. U., & Emina, J. O. (2009). Stall in fertility decline in Eastern African countries: Regional analysis of patterns, determinants and implications. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1532), 2991–3007. https://doi.org/10.1098/rstb.2009.0166

Fatehkia, M., Kashyap, R., & Weber, I. (2018). Using Facebook ad data to track the global digital gender gap. *World Development*, *107*, 189–209. https://doi.org/10.1016/j.worlddev.2018.03.007

Ferrara, E. L., Chong, A., & Duryea, S. (2012). Soap operas and fertility: Evidence from Brazil. *American Economic Journal: Applied Economics*, *4*(4), 1–31.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140.

Feyissa, G. T., Tolu, L. B., Soboka, M., & Ezeh, A. (2023). Effectiveness of interventions to reduce child marriage and teen pregnancy in sub-Saharan Africa: A systematic review of quantitative evidence. *Frontiers in Reproductive Health*, *5*, 1105390.

Field, E., & Ambrus, A. (2008). Early marriage, age of menarche, and female schooling attainment in Bangladesh. *Journal of Political Economy*, *116*(5), 881–930.

Fischer, R. (2009). Where is culture in cross cultural research? An outline of a multilevel research process for measuring culture as a shared meaning system. *International Journal of Cross Cultural Management*, *9*(1), 25–49.

Flückiger, M., & Ludwig, M. (2023). Mobile phone coverage and infant mortality in sub-Saharan Africa. *Journal of Economic Behavior & Organization*, *211*, 462–485.

Ford, M. E., & Kelly, P. A. (2005). Conceptualizing and categorizing race and ethnicity in health services research. *Health Services Research*, *40*(5p2), 1658–1675.

Ganchimeg, T., Ota, E., Morisaki, N., Laopaiboon, M., Lumbiganon, P., Zhang, J., Yamdamsuren, B., Temmerman, M., Say, L., Tunçalp, Ö., & others. (2014). Pregnancy and childbirth outcomes among adolescent mothers: A W orld H ealth O rganization multicountry study. *BJOG: An International Journal of Obstetrics & Gynaecology*, *121*, 40–48.

Garenne, M. (2004). Age at marriage and modernisation in sub-Saharan Africa. *Southern African Journal of Demography*, *9*(2), 59–79.

Garenne, M. (2022). Marriage in sub-saharan africa. In *The Routledge Handbook of African Demography* (pp. 151–180). Routledge.

Geldof, M., Thiombiano, B. A., & Wagner, N. (2021). "When I receive the message, it is a sign of love": Symbolic connotations of SMS messages for people living with HIV in Burkina Faso. *AIDS Care*, *33*(6), 810–817. https://doi.org/10.1080/09540121.2020.1769832

Gelfand, M. J., Nishii, L. H., & Raver, J. L. (2006). On the nature and importance of cultural tightness-looseness. *Journal of Applied Psychology*, *91*(6), 1225.

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliach, A., Ang, S., Arnadottir, J., & others. (2011). Differences between tight and loose cultures: A 33-nation study. *Science*, *332*(6033), 1100–1104.

Gerland, P., Biddlecom, A., & Kantorová, V. (2017). Patterns of Fertility Decline and the Impact of Alternative Scenarios of Future Fertility Change in sub-Saharan Africa. *Population and Development Review*, *43*(S1), 21–38. https://doi.org/10.1111/padr.12011

Gittleman, J. L., & Kot, M. (1990). Adaptation: Statistics and a null model for estimating phylogenetic effects. *Systematic Zoology*, *39*(3), 227–241.

Goldin, C. (2025). *The Downside of Fertility* (Working Paper No. 34268). National Bureau of Economic Research. https://doi.org/10.3386/w34268

Goldscheider, C., & Uhlenberg, P. R. (1969). Minority group status and fertility. *American Journal of Sociology*, *74*(4), 361–372.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, *91*(3), 481–510. https://doi.org/10.1086/228311

Guldi, M., & Herbst, C. M. (2017). Offline effects of online connecting: The impact of broadband diffusion on teen fertility decisions. *Journal of Population Economics*, *30*, 69–91.

Gündüz-Hoşgör, A., & Smits, J. (2008). Variation in labor market participation of married women in Turkey. *Women's Studies International Forum*, *31*(2), 104–117. https://doi.org/10.1016/J.WSIF.2008.03.003

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

Guriev, S., Melnikov, N., & Zhuravskaya, E. (2021). 3g internet and confidence in government. *Quarterly Journal of Economics*, *136*(4), 2533–2613.

Gurumurthy, A., & Chami, N. (2014). Gender equality in the information society. *Bangalore: IT for Change*.

Henrich, J. (2015). The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter. In *The secret of our success*. princeton University press.

Hertrich, V. (2017). Trends in Age at Marriage and the Onset of Fertility Transition in sub-Saharan Africa. *Population and Development Review*, *43*(S1), 112–137. https://doi.org/10.1111/padr.12043

Hicks, J., & Hicks, D. L. (2019). Lucky Late Bloomers? Consequences of Delayed Marriage for Women in Rural Western Kenya. *SSRN No. 2622189. \normalfontAvailable at SSRN: Http://Dx.Doi.Org/10.2139/Ssrn.2622189*.

Hindin, M. J. (2000). Women's autonomy, women's status and fertility-related behavior in Zimbabwe. *Population Research and Policy Review*, *19*(3), 255–282. https://doi.org/10.1023/A:1026590717779

Hutcheon, J. A., Chiolero, A., & Hanley, J. A. (2010). Random measurement error and regression dilution bias. *BMJ*, *340*, c2289. https://doi.org/10.1136/bmj.c2289

ICRW. (2007). *New insights on preventing child marriage. A global analysis of factors and programs*. International Center for Research on Women.

Inkeles, A., & Smith, D. H. (1974). *Becoming modern: Individual change in six developing countries*. Harvard University Press.

Inkeles, A., & Smith, D. H. (1985). Becoming modern. In *The Gap Between Rich And Poor* (pp. 70–77). Routledge.

International Telecom Union. (2024). *ITU Digital development: ICT indicator database*.

Jensen, R. (2007). The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector. *Quarterly Journal of Economics*, *122*(3), 879–924.

Jensen, R., & Oster, E. (2009). The power of TV: Cable television and women's status in India. *The Quarterly Journal of Economics*, *124*(3), 1057–1094.

Jensen, R., & Thornton, R. (2003). Early female marriage in the developing world. *Gender & Development*, *11*(2), 9–19.

Johnson-Hanks, J. (2005). When the Future Decides: Uncertainty and Intentional Action in Contemporary Cameroon. *Current Anthropology*, *46*(3), 363–385. https://doi.org/10.1086/428799

Johnson-Hanks, J. (2006). *Uncertain Honor: Modern Motherhood in an African Crisis*. University of Chicago Press.

Johnson-Hanks, J. (2007). Natural Intentions: Fertility Decline in the African Demographic and Health Surveys. *American Journal of Sociology*, *112*(4), 1008–1043. https://doi.org/10.1086/508791

Johnson-Hanks, J. A., Bachrach, C. A., Morgan, S. P., & Kohler, H.-P. (2011). *Understanding family change and variation: Toward a theory of conjunctural action* (Vol. 5). Springer Science & Business Media.

Kane, S., Kok, M., Rial, M., Matere, A., Dieleman, M., & Broerse, J. E. (2016). Social norms and family planning decisions in South Sudan. *BMC Public Health*, *16*(1), 1183. https://doi.org/10.1186/s12889-016-3839-6

Kashyap, R., Fatehkia, M., Tamime, R. A., & Weber, I. (2020). Monitoring global digital gender inequality using the online populations of Facebook and Google. *Demographic Research*, *43*, 779–816. https://doi.org/10.4054/DemRes.2020.43.27

Kassa, G. M., Arowojolu, A., Odukogbe, A., & Yalew, A. W. (2018). Prevalence and determinants of adolescent pregnancy in Africa: A systematic review and meta-analysis. *Reproductive Health*, *15*, 1–17.

Kearney, M. S., & Levine, P. B. (2015). Media influences on social outcomes: The impact of MTV's 16 and pregnant on teen childbearing. *American Economic Review*, *105*(12), 3597–3632.

Kebede, E., Goujon, A., & Lutz, W. (2019). Stalls in Africa's fertility decline partly result from disruptions in female education. *Proceedings of the National Academy of Sciences*, *116*(8), 2891–2896. https://doi.org/10.1073/pnas.1717288116

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, *54*(3), 241–251. https://doi.org/10.1016/j.bushor.2011.01.005

Kolk, M. (2014). Understanding transmission of fertility across multiple generations – Socialization or socioeconomics? *Research in Social Stratification and Mobility*, *35*, 89–103. https://doi.org/10.1016/j.rssm.2013.09.006

Kollehlon, K. T. (1989). Ethnicity and fertility in Liberia: A test of the minority group status hypothesis. *Social Biology*, *36*(1–2), 67–81. https://doi.org/10.1080/19485565.1989.9988720

Kollehlon, K. T. (2003). Ethnicity and fertility in Nigeria. *Social Biology*, *50*(3–4), 201–221. https://doi.org/10.1080/19485565.2003.9989072

Koski, A., Clark, S., & Nandi, A. (2017). Has child marriage declined in sub-Saharan Africa? An analysis of trends in 31 countries. *Population and Development Review*, 7–29.

Kravdal, Ø. (2006). A simulation-based assessment of the bias produced when using averages from small DHS clusters as contextual variables in multilevel models. *Demographic Research*, *15*, 1–20. https://doi.org/10.4054/DemRes.2006.15.1

Kravdal, Ø. (2012). Further evidence of community education effects on fertility in sub-Saharan Africa. *Demographic Research*, *27*, 645–680. https://doi.org/10.4054/DemRes.2012.27.22

Kummu, M., Taka, M., & Guillaume, J. H. A. (2018). Gridded global datasets for Gross Domestic Product and Human Development Index over 1990–2015. *Scientific Data*, *5*(1), 180004. https://doi.org/10.1038/sdata.2018.4

Lembani, R., Gunter, A., Breines, M., & Dalu, M. T. B. (2020). The same course, different access: The digital divide between urban and rural distance education students in South Africa. *Journal of Geography in Higher Education*, *44*(1), 70–84.

Lerch, M. (2019). Fertility Decline in Urban and Rural Areas of Developing Countries. *Population and Development Review*, *45*(2), 301–320. https://doi.org/10.1111/padr.12220

Lerner, D. (1958). *The passing of traditional society: Modernizing the Middle East*.

Lesthaeghe, R. J. (1978). *The decline of Belgian fertility, 1800-1970*. Princeton University Press.

Levy, D. A., & Nail, P. R. (1993). Contagion: A theoretical and empirical review and reconceptualization. *Genetic, Social, and General Psychology Monographs*, *119*(2), 233–284.

Little, R. J., & Rubin, D. B. (2019). *Statistical analysis with missing data*. John Wiley & Sons.

Lois, D., & Arránz Becker, O. (2014). Is fertility contagious? Using panel data to disentangle mechanisms of social network influences on fertility decisions. *Advances in Life Course Research*, *21*, 123–134. https://doi.org/10.1016/j.alcr.2013.10.001

Longwe, A., & Smits, J. (2012). Family Planning Outcomes and Primary School Attendance in Sub-Saharan Africa. *Studies in Family Planning*, *43*(2), 127–134. https://doi.org/10.1111/j.1728-4465.2012.00310.x

Longwe, A., & Smits, J. (2013). The impact of family planning on primary school enrolment in sub-national areas within 25 African countries. *African Journal of Reproductive Health*, *17*(2), 23–38.

Manacorda, M., & Tesei, A. (2020). Liberation technology: Mobile phones and political mobilization in Africa. *Econometrica*, *88*(2), 533–567.

Manser, M., & Brown, M. (1980). Marriage and Household Decision-Making: A Bargaining Analysis. *International Economic Review*, *21*(1), 31–44.

Maralani, V. (2008). The changing relationship between family size and educational attainment over the course of socioeconomic development: Evidence from Indonesia. *Demography*, *45*(3), 693–717. https://doi.org/10.1353/dem.0.0013

Matovu, J. K. B., Makumbi, F., Wanyenze, R. K., & Serwadda, D. (2017). Determinants of fertility desire among married or cohabiting individuals in Rakai, Uganda: A cross-sectional study. *Reproductive Health*, *14*(1), 2. https://doi.org/10.1186/s12978-016-0272-3

Mbaye, L. M., & Wagner, N. (2017). Bride Price and Fertility Decisions: Evidence from Rural Senegal. *The Journal of Development Studies*, *53*(6), 891–910. https://doi.org/10.1080/00220388.2016.1208178

McElroy, M. B., & Horney, M. J. (1981). Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Demand. *International Economic Review*, *22*(2), 333–349.

Menon, J., Kusanthan, T., Mwaba, S., Juanola, L., & Kok, M. (2018). 'Ring'your future, without changing diaper–Can preventing teenage pregnancy address child marriage in Zambia? *PloS One*, *13*(10), e0205523.

Meta. (2019). *An Update on How We Are Doing At Enforcing Our Community Standards*. https://about.fb.com/news/2019/05/enforcing-our-community-standards-3/

Mkhwanazi, N. (2010). Understanding teenage pregnancy in a post-apartheid South African township. *Culture, Health & Sexuality*, *12*(4), 347–358.

Moursund, A., & Kravdal, Ø. (2003). Individual and community effects of women's education and autonomy on contraceptive use in India. *Population Studies*, *57*(3), 285–301. https://doi.org/10.1080/0032472032000137817

Munshi, K., & Myaux, J. (2006). Social norms and the fertility transition. *Journal of Development Economics*, *80*(1), 1–38. https://doi.org/10.1016/j.jdeveco.2005.01.002

Nanfuka, J. (2022). Digital Access and Economic Transformation in Africa. *Institute for New Economic Thinking. Africa Paper Series*, *4*.

Nkoa, B. E. O., Beyene, B. O., Simb, J. F. N. N., & Eloundou, G. N. (2023). Does social media improve women's political empowerment in Africa? *Telecommunications Policy*, *47*(9), 102624.

Norris, P. (2001). Digital divide: Civic engagement, information poverty, and the Internet worldwide. *Cambridge University Press Google Schola*, *2*, 12–136.

Ofori, K. N., & Mohangi, K. (2024). Adolescent psychosocial identity development associated with traditional rituals and rites of passage in the Bosomtwe District of Ghana. *Vulnerable Children and Youth Studies*, *19*(4), 733–750. https://doi.org/10.1080/17450128.2024.2413528

Okenwa-Emegwa, L., Lawoko, S., & Jansson, B. (2016). Attitudes Toward Physical Intimate Partner Violence Against Women in Nigeria. *SAGE Open*, *6*(4), 2158244016667993. https://doi.org/10.1177/2158244016667993

Ord, J. K., & Getis, A. (1995). Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis*, *27*(4), 286–306. https://doi.org/10.1111/j.1538-4632.1995.tb00912.x

Palamuleni, M. E. (2014). Social and economic factors affecting ethnic fertility differentials in Malawi. *International Journal of Development and Sustainability*, *3*(1), 70–88.

Petersen, W. (1969). *Population*. Macmillan.

Pezzulo, C., Nilsen, K., Carioli, A., Tejedor-Garavito, N., Hanspal, S. E., Hilber, T., James, W. H. M., Ruktanonchai, C. W., Alegana, V., Sorichetta, A., Wigley, A. S., Hornby, G. M., Matthews, Z., & Tatem, A. J. (2021). Geographical distribution of fertility rates in 70 low-income, lower-middle-income, and upper-middle-income countries, 2010–16: A subnational analysis of cross-sectional surveys. *The Lancet Global Health*, *9*(6), e802–e812. https://doi.org/10.1016/S2214-109X(21)00082-6

Pierotti, R. S. (2013). Increasing Rejection of Intimate Partner Violence: Evidence of Global Cultural Diffusion. *American Sociological Review*, *78*(2), 240–265. https://doi.org/10.1177/0003122413480363

Porter, G., Hampshire, K., Abane, A., Munthali, A., Robson, E., Bango, A., de Lannoy, A., Gunguluza, N., Tanle, A., Owusu, S., & others. (2015). Intergenerational relations and the power of the cell phone: Perspectives on young people's phone usage in sub-Saharan Africa. *Geoforum*, *64*, 37–46.

Porter, G., Hampshire, K., Abane, A., Munthali, A., Robson, E., De Lannoy, A., Tanle, A., & Owusu, S. (2020). Mobile phones, gender, and female empowerment in sub-Saharan Africa: Studies with African youth. *Information Technology for Development*, *26*(1), 180–193.

Preston-Whyte, E. M., & Zondi, M. (1991). Adolescent sexuality and its implications for teenage pregnancy and AIDS. *CME: South Africa*, *9*(11), 1389–1394.

Putnam, R. D. (1993). The prosperous community. *The American Prospect*, *4*(13), 35–42.

Rainie, L., & Wellman, B. (2012). *Networked: The New Social Operating System*. MIT Press. https://doi.org/10.7551/mitpress/8355.001.0001

Rampazzo, F., Bijak, J., Vitali, A., Weber, I., & Zagheni, E. (2021). A Framework for Estimating Migrant Stocks Using Digital Traces and Survey Data: An Application in the United Kingdom. *Demography*, *58*(6), 2193–2218. https://doi.org/10.1215/00703370-9578562

Ribeiro, F. N., Benevenuto, F., & Zagheni, E. (2020). How Biased is the Population of Facebook Users? Comparing the Demographics of Facebook Users with Census Data to Generate Correction Factors. *Proceedings of the 12th ACM Conference on Web Science*, 325–334. https://doi.org/10.1145/3394231.3397923

Robinson, A. L. (2014). National versus Ethnic Identification in Africa: Modernization, Colonial Legacy, and the Origins of Territorial Nationalism. *World Politics*, *66*(4), 709–746.

Robinson, L., Cotten, S. R., Ono, H., Quan-Haase, A., Mesch, G., Chen, W., Schulz, J., Hale, T. M., & Stern, M. J. (2015). Digital inequalities and why they matter. *Information, Communication & Society*, *18*(5), 569–582. https://doi.org/10.1080/1369118X.2015.1012532

Rotondi, V., Kashyap, R., Pesando, L. M., Spinelli, S., & Billari, F. C. (2020). Leveraging mobile phones to attain sustainable development. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(24), 13413–13420. https://doi.org/10.1073/pnas.1909326117

Schaffnit, S. B., Urassa, M., & Lawson, D. W. (2019). "Child marriage" in context: Exploring local attitudes towards early marriage in rural Tanzania. *Sexual and Reproductive Health Matters*, *27*(1), 93–105.

Scheerder, A., van Deursen, A., & van Dijk, J. (2017). Determinants of Internet skills, uses and outcomes. A systematic review of the second- and third-level digital divide. *Telematics and Informatics*, *34*(8), 1607–1624. https://doi.org/10.1016/j.tele.2017.07.007

Schrijner, S., & Smits, J. (2018). Grandmothers and Children's Schooling in Sub-Saharan Africa. *Human Nature*, *29*(1), 65–89. https://doi.org/10.1007/s12110-017-9306-y

Schroeder, E., Tallarico, R., & and, M. B. (2022). The impact of adolescent initiation rites in East and Southern Africa: Implications for policies and practices. *International Journal of Adolescence and Youth*, *27*(1), 181–192. https://doi.org/10.1080/02673843.2022.2052123

Sedgh, G., & Hussain, R. (2014). Reasons for contraceptive nonuse among women having unmet need for contraception in developing countries. *Studies in Family Planning*, *45*(2), 151–169. https://doi.org/10.1111/j.1728-4465.2014.00382.x

Sennott, C., & Yeatman, S. (2012). Stability and change in fertility preferences among young women in Malawi. *International Perspectives on Sexual and Reproductive Health*, *38*(1), 34.

Shapiro, D., & Gebreselassie, T. (2014). Marriage in sub-Saharan Africa: Trends, determinants, and consequences. *Population Research and Policy Review*, *33*, 229–255.

Shapiro, D., & Tambashe, B. O. (1997). Education, employment, and fertility in Kinshasa and prospects for changes in reproductive behavior. *Population Research and Policy Review*, *16*(3), 259–287. https://doi.org/10.1023/A:1005761504449

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*, 669–708. https://doi.org/10.4054/DemRes.2017.37.21

Smits, J., & Permanyer, I. (2019). The Subnational Human Development Database. *Scientific Data*, *6*(1), 190038. https://doi.org/10.1038/sdata.2019.38

Smits, J., & Permanyer, I. (2020). *The Subnational Gender Development Index: Within-country variation in gender (in) equality revealed*. UNDP HDialogue Blog. https://globaldatalab. org/blog-SGDI.

Smits, J., & Steendijk, R. (2015). The International Wealth Index (IWI). *Social Indicators Research*, *122*(1), 65–85. https://doi.org/10.1007/s11205-014-0683-x

Snijders, T. A. B., & Bosker, R. (2011). Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. *Sage Publiciations, Inc.*, 1–368.

Spierings, N., Smits, J., & Verloo, M. (2010). Micro-and Macrolevel Determinants of Women's Employment in Six Arab Countries. *Journal of Marriage and Family*, *72*(5), 1391–1407. https://doi.org/10.1111/j.l741-3737.2010.00772.x

Sripad, P., Pinchoff, J., Dadi, C., & Dougherty, L. (2024). Measuring social norms related to child marriage among married women and men in Niger. *PLOS ONE*, *19*(7), 1–16. https://doi.org/10.1371/journal.pone.0307595

Stevanovic-Fenn, N., Edmeades, J., Lantos, H., & Onovo, O. (2015). Child marriage, adolescent pregnancy and family formation in West and Central Africa: Patterns, trends, and drivers of change. *Dakar, Senegal: UNICEF West and Central Africa Regional Office*.

Stewart, I., Flores, R. D., Riffe, T., Weber, I., & Zagheni, E. (2019). Rock, Rap, or Reggaeton?: Assessing Mexican Immigrants' Cultural Assimilation Using Facebook Data,. *The World Wide Web Conference*, 3258–3264. https://doi.org/10.1145/3308558.3313409

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199.

Tewahido, D., Worku, A., Tadesse, A. W., Gulema, H., & Berhane, Y. (2022). Adolescent girls trapped in early marriage social norm in rural Ethiopia: A vignette-based qualitative exploration. *PLOS ONE*, *17*(2), 1–14. https://doi.org/10.1371/journal.pone.0263987

Toffolutti, V., Ma, H., Menichelli, G., Berlot, E., Mencarini, L., & Aassve, A. (2020). How the internet increases modern contraception uptake: Evidence from eight sub-Saharan African countries. *BMJ Global Health*, *5*(11), e002616.

Tomkinson, J. (2019). Age at first birth and subsequent fertility: The case of adolescent mothers in France and England and Wales. *Demographic Research*, *40*, 761–798. https://doi.org/10.4054/DemRes.2019.40.27

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

UNESCO. (2021). *The journey towards comprehensive sexuality education: Global status report*. https://www.unesco.org/en/articles/journey-towards-comprehensive-sexuality-education-global-status-report

UNFPA. (2012). *Marrying too young: End child marriage*. New York: UNFPA.

UNICEF. (2015). *Child marriage, Adolescent pregnancy and Family formation in West and Central Africa*. United Nations Children's Fund report.

UNICEF. (2023). *Bridging the gender digital divide: Challenges and an Urgent Call for Action for Equitable Digital Skills Development*. United Nations Children's Fund report. https://data.unicef.org/resources/ictgenderdivide/

UNICEF. (2024). *Early childbearing*.

United Nations. (2020). *World fertility and family planning 2020: Highlights*. United Nations. https://www.un.org/development/desa/pd/content/world-fertility-and-family-planning-2020-highlights

United Nations. (2023). *World Population Prospects 2023: Summary of Results*. United Nations. https://population.un.org/wpp/

Upadhyay, U. D., Gipson, J. D., Withers, M., Lewis, S., Ciaraldi, E. J., Fraser, A., Huchko, M. J., & Prata, N. (2014). Women's empowerment and fertility: A review of the literature. *Social Science & Medicine*, *115*, 111–120. https://doi.org/10.1016/j.socscimed.2014.06.014

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*, 78–89.

Varga, C. A. (2003). How Gender Roles Influence Sexual and Reproductive Health Among South African Adolescents. *Studies in Family Planning*, *34*(3), 160–172. https://doi.org/10.1111/j.1728-4465.2003.00160.x

Varriale, C., Pesando, L. M., Kashyap, R., & Rotondi, V. (2022). Mobile Phones and Attitudes toward Women's Participation in PoliticsEvidence from Africa. *Sociology of Development*, *8*(1), 1–37. https://doi.org/10.1525/sod.2020.0039

Viollaz, M., & Winkler, H. (2022). Does the internet reduce gender gaps? The case of Jordan. *The Journal of Development Studies*, *58*(3), 436–453.

Wagner, N. (2015). Female Genital Cutting and Long-Term Health Consequences ‚Äì Nationally Representative Estimates across 13 Countries. *The Journal of Development Studies*, *51*(3), 226–246. https://doi.org/10.1080/00220388.2014.976620

Westoff, C. F., Bankole, A., & others. (1997). *Mass media and reproductive behavior in Africa*. Macro International Calverton, MD.

Westoff, C. F., & Cross, A. R. (2006). *The stall in the fertility transition in Kenya*.

Westoff, C. F., & Koffman, D. A. (2011). The association of television and radio with reproductive behavior. *Population and Development Review*, *37*(4), 749–759.

WHO. (2024). *Adolescent pregnancies*.

Wildeman, J., Schrijner, S., & Smits, J. (2023). Fertility rates and social media usage in sub-Saharan Africa. *Population, Space and Place*, *29*(4), e35.

Wilkinson, O., McDonald, K., Trotta, S., Eggert, J. P., & de Wolf, F. (2024). A Social Norms Analysis of Religious Drivers of Child Marriage. *Global Health: Science and Practice*, *12*(2). https://doi.org/10.9745/GHSP-D-23-00339

Wing, C., Freedman, S. M., & Alex, H. (2024). *Stacked difference-in-differences* (Working Paper No. 32054; Working Paper Series). National Bureau of Economic Research. http://www.nber.org/papers/w32054

Withers, M., Dworkin, S. L., Onono, M., Oyier, B., Cohen, C. R., Bukusi, E. A., & Newmann, S. J. (2015). Men's Perspectives on Their Role in Family Planning in Nyanza Province, Kenya. *Studies in Family Planning*, *46*(2), 201–215. https://doi.org/10.1111/j.1728-4465.2015.00024.x

World Bank. (2021a). *Birth rate, crude (per 1,000 people)*. https://data.worldbank.org/indicator/SP.DYN.CBRT.IN

World Bank. (2021b). *Mobile cellular subscriptions (per 1000 people)*. https://data.worldbank.org/indicator/IT.CEL.SETS.P2

Yakubu, I., & Salisu, W. J. (2018). Determinants of adolescent pregnancy in sub-Saharan Africa: A systematic review. *Reproductive Health*, *15*, 1–11.

Zagheni, E., Weber, I., & Gummadi, K. (2017). Leveraging Facebook's Advertising Platform to Monitor Stocks of Migrants. *Population and Development Review*, *43*(4), 721–734. https://doi.org/10.1111/padr.12102

Zgambo, T. (2022). The construction of teen pregnancy and child marriage in africa. In *The Routledge Handbook of African Demography* (pp. 181–197). Routledge.

Appendices

Summary in English

This thesis reveals the importance of social contexts and the role of digital infrastructure in shaping women's timing of family formation and family size preferences in SSA. Using large-scale demographic and geospatial data, the analyses examine variation in family size preferences, the timing of marriage and first birth, and the prevalence of child marriage and adolescent childbearing.

Across the African region, substantial differences in family formation patterns are observed between ethnic, religious, and local community groups. Ethnic group membership is associated with marked variation in preferred family size, even after adjusting for women's education, household wealth, and other individual factors. Group-level characteristics such as average education and gender-related norms further explain these differences, and their influence varies depending on a woman's own education and occupation.

A novel measure of normative strength, based on the degree of variation in marriage and childbirth timing within social groups, shows that lower variation —interpreted as stronger group norms—is linked to earlier transitions into marriage and motherhood, and to higher rates of child marriage and teen pregnancy. This association is most pronounced when the reference group consists of age-similar peers, indicating that peer norms are particularly influential for fertility timing.

Digital connectivity indicators reveal systematic associations with reproductive outcomes. Regions with higher rates of social media use and narrower gender gaps in digital access exhibit lower crude birth rates, controlling for socioeconomic conditions. Similarly, the introduction of mobile broadband (3G) coverage is followed by significant reductions in child marriage and teen pregnancy, with the largest effects among more educated and wealthier adolescent girls. Exploratory analyses suggest that increased access to information, changes in gender attitudes, and behavioral adaptations such as contraceptive use are likely mechanisms.

Taken together, the results document wide heterogeneity in women's reproductive preferences and behaviors across SSA, quantify the role of social contexts in shaping these outcomes, and show that digital infrastructure is associated with shifts toward later and less frequent childbearing in certain segments of the population.

Nederlandse samenvatting

Dit proefschrift laat zien hoe belangrijk de sociale context en digitale infrastructuur zijn voor de timing van gezinsvorming en de voorkeuren voor gezinsgrootte van vrouwen in Sub-Sahara Afrika (SSA). Met behulp van grootschalige demografische en geografische data zijn variaties onderzocht in voorkeuren voor gezinsgrootte, de timing van het huwelijk en de eerste geboorte, en de prevalentie van kindhuwelijken en tienerzwangerschappen.

Binnen SSA zijn aanzienlijke verschillen in gezinsvormingspatronen gevonden tussen etnische, religieuze en lokale gemeenschappen. Zo blijkt dat het lidmaatschap van een etnische groep sterk samenhangt met verschillen in de gewenste gezinsgrootte, zelfs na correctie voor opleiding, huishoudelijk welvaartsniveau en andere individuele kenmerken van vrouwen. Daarnaast bleken groepskenmerken, zoals gemiddeld opleidingsniveau en gendergerelateerde normen, deze verschillen verder te verklaren. De mate van groepsinvloed hangt bovendien sterk af van het individuele opleidingsniveau en beroep van de vrouw.

Een nieuwe indicator voor normatieve sterkte, gebaseerd op de mate van variatie in de timing van huwelijk en geboorte binnen sociale groepen, laat zien dat een lage variatie – geïnterpreteerd als een sterkere groepsnorm – samenhangt met eerdere overgangen naar huwelijk en moederschap en met hogere percentages kindhuwelijken en tienerzwangerschappen. Deze samenhang is het sterkst wanneer de referentiegroep bestaat uit leeftijdsgenoten, wat erop wijst dat normen die door leeftijdsgenoten worden gezet bijzonder invloedrijk zijn voor de timing van vruchtbaarheid.

Er is een systematisch verband tussen indicatoren van digitale connectiviteit en reproductieve uitkomsten. Regio's met hoger sociale-mediagebruik en kleinere genderverschillen in digitale toegang vertonen lagere geboortecijfers, zelfs na controle voor sociaaleconomische omstandigheden. De resultaten laten ook zien dat de komst van mobiele breedband (3G)-internet leidt tot significante dalingen in het aantal kindhuwelijken en tienerzwangerschappen, met de grootste effecten onder hoger opgeleide en welvarendere jonge vrouwen. Verdere analyses suggereren dat meer toegang tot informatie, veranderende genderhoudingen en gedragsaanpassingen, zoals het gebruik van anticonceptie, waarschijnlijk de onderliggende mechanismen zijn.

Gezamenlijk tonen de resultaten aan dat er binnen Sub-Sahara Afrika grote heterogeniteit bestaat in de reproductieve voorkeuren en gedragingen van vrouwen, dat de normen van de sociale context een belangrijke rol spelen in deze uitkomsten, en dat digitale infrastructuur samenhangt met verschuivingen naar later en minder frequent kinderen krijgen en trouwen.

Research data management

The significance of data management in maintaining the quality of scientific work has grown considerably. Consequently, Radboud University has established stringent criteria for researchers to follow when storing scientific data and facilitating the replication of studies by others. In this section, I outline the measures implemented to safeguard the quality of my data.

This dissertation used secondary data obtained from various publicly accessible sources; no primary data were collected for this research. Major providers of open-access datasets include the Demographic and Health Surveys (DHS), the Multiple Indicator Cluster Surveys (MICS), the World Bank, Collins Bartholomew, the Global Data Lab, and Afrobarometer, among others. The datasets, along with syntax files (in .sav and .rds formats), are stored in Radboud University's workgroup folders. For data cleaning, analysis, and reporting, I used the statistical software SPSS and R. For each chapter, a syntax file was created, detailing the modifications and empirical analyses. The data management procedures implemented in this dissertation comply with the regulatory frameworks and guidelines set forth by Radboud University.

Author contributions

Chapter 1:
Conceptualization: Wildeman
Investigation: Wildeman
Writing – Original Draft: Wildeman
Writing – Review & Editing: Wildeman, Smits, and Wagner

Chapter 2:
Conceptualization: Wildeman, Schrijner, and Smits
Methodology: Wildeman, Schrijner, and Smits
Formal analysis: Wildeman
Investigation: Wildeman
Data curation: Wildeman
Writing – original draft: Wildeman
Writing – review and editing: Wildeman, Schrijner, and Smits

Chapter 3:
Conceptualization: Wildeman and Wagner
Methodology: Wildeman and Wagner
Formal analysis: Wildeman
Investigation: Wildeman
Data curation: Wildeman
Writing – original draft: Wildeman
Writing – review and editing: Wildeman and Wagner

Chapter 4:
Conceptualization: Wildeman, Schrijner, and Smits
Methodology: Wildeman, Schrijner, and Smits
Formal analysis: Wildeman
Investigation: Wildeman
Data curation: Wildeman
Writing – original draft: Wildeman
Writing – review and editing: Wildeman, Schrijner, and Smits

Chapter 5:
Conceptualization: Wildeman and Wagner
Methodology: Wildeman, Morsink and Wagner
Formal analysis: Wildeman
Investigation: Wildeman
Data curation: Wildeman
Writing – original draft: Wildeman, Morsink and Wagner
Writing – review and editing: Wildeman, Morsink and Wagner

Chapter 6:
Conceptualization: Wildeman
Writing – original draft: Wildeman
Writing – review & editing: Wildeman, Smits and Wagner

Dankwoord

Aan het einde van mijn promotie traject is het tijd om mijn dank uit te spreken. Deze dissertatie was er nooit gekomen zonder de steun van velen. Ik wil graag iedereen bedanken die mij in de afgelopen jaren heeft geholpen. Elk bijdrage, groot of klein, wordt gewaardeerd.

In het bijzonder wil ik mijn promotors bedanken. Jeroen, in een keuzevak tijdens mijn master leerde ik het vakgebied ontwikkelingseconomie en het Global Data Lab kennen. Hierdoor leerde ik dat er een vakgebied was waarin ik mijn interesses in niet-Westerse landen en kwantitatieve data analyse kon combineren. Ik had het geluk dat er een PhD positie vrijkwam en dat ik aangenomen werd en zo begon onze samenwerking. Jouw harde werk voor het Global Data Lab, je betrokkenheid bij het onderzoek en je enthousiasme zijn een grote inspiratie voor mij.

Natascha, jij kwam halverwege mijn PhD bij mijn project en je leerde mij niet alleen veel over het economische aspect van het vakgebied, maar je liet mij ook zien hoe je als moeder een uitdagende baan kan combineren met een gezin. We waren zelfs tegelijk zwanger, hoe toevallig. Daarnaast heb je mij enorm geholpen en vertrouwen gegeven om meer de economische richting uit te gaan in mijn onderzoek. Bedankt voor alle steun en ook alle humor tijdens onze besprekingen.

Dit proefschrift was er ook niet geweest zonder de hulp van Sandor Schrijner en Karlijn Morsink. Sandor, in het begin van mijn PhD was jij betrokken bij mijn project en je bracht veel gezelligheid en sfeer bij de meetings voor mijn eerste artikel en de eerste 'officiële' GDL vergaderingen. Karlijn, wij leerden elkaar kennen bij een workshop van ontwikkelingseconomie en niet veel later begon onze samenwerking. Inhoudelijke discussies over wat de beste methode is en hard werken volgden elkaar op met de nodige humor. We kunnen heel tevreden zijn over het resultaat, waarvan ik het vertrouwen heb dat het resulteert in een mooie publicatie.

Ik wil ook graag al mijn collega promovendi in Nijmegen bedanken voor de steun die ze me afgelopen jaren hebben gegeven. In het bijzonder wil ik mijn 'oude' kantoorgenoten bedanken, Robbert, Charan, Rutger en Annemiek, voor jullie gezelligheid en de vele koffiepauze. Maar ook voor inhoudelijk advies en hulp bij het geven van mijn eerste onderwijs. Daarnaast wil ik ook Anne Lieke en Emily bedanken, als collega PhD's binnen hetzelfde vakgebied. Helaas is het ons nooit gelukt om naar dezelfde conferentie te gaan, maar ik heb enorm veel aan jullie gehad. Tot slot wil ik ook Simon, Aaron, Roben, Siem en Oskar bedanken voor de ontspanning en gezelligheid!

Ten slotte wil ik nog mijn familie bedanken voor alle steun. Lieve Thijmen, we begonnen tegelijk aan een PhD en hierdoor begrepen we elkaar heel goed. Je was zelfs even mijn favoriete (en enige) thuis-kantoorgenoot tijdens corona. Dit hebben we voortreffelijk doorstaan, al zeg ik het zelf, en ik kan nog steeds altijd op je rekenen. Dankjewel. En dan natuurlijk de komst van Mart en Tibbe niet te vergeten. Hoewel het hierdoor soms uitdagender was om te werken, hielpen zij juist met prioriteren en efficiënt werken. En natuurlijk hielpen alle lieve knuffels en kusjes ook ter motivatie.

About the author

Jet Wildeman was born on December 7th, 1993, in Raalte, the Netherlands. She obtained her bachelor's degree in psychology from Radboud University in 2016, followed by a Research master's degree in social and Cultural Science from the same university in 2019. She then joined the Department of Economics at Radboud University as a PhD candidate, focusing on the social, economic, and cultural determinants of marriage and fertility in sub-Saharan Africa. Her research combines large-scale survey data, satellite imagery, and digital trace data to study how local contexts and digital connectivity shape demographic behavior. During her doctoral studies, she also worked as a data specialist at the Global Data Lab, contributing to data infrastructure and visualization. Jet has presented her work at international conferences, published in peer-reviewed journals, and won the Best Poster Award at the Population Association of America Conference in 2021. She is fluent in Dutch and English and has a strong interest in making research results accessible to a broad audience.

Radboud Universiteit