



Infrastructures, population dynamics and internal migrations in Sub-Saharan Africa

Hamidou Diallo

► To cite this version:

Hamidou Diallo. Infrastructures, population dynamics and internal migrations in Sub-Saharan Africa. Economics and Finance. Université Paris sciences et lettres, 2022. English. NNT : 2022UPSLD028 . tel-04051824

HAL Id: tel-04051824

<https://theses.hal.science/tel-04051824>

Submitted on 30 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'Université Paris Dauphine - PSL

**Infrastructures, Population dynamics, and Internal
Migrations in Sub-Saharan Africa**

Soutenue par

Hamidou DIALLO

Le 20 Décembre 2022

École doctorale n°543

SDOSE

Spécialité

Sciences Economiques

Composition du jury :

Isabelle CHORT

Professeur, Université de Pau et des
pays de l'Adour *Rapporteuse*

Philippe DE VREYER

Professeur, Université Paris Dauphine *Président du jury*

Karine MARAZYAN

Professeur, Université de Rouen *Rapporteuse*

Sandrine MESPLÉ-SOMPS

Directrice de Recherche, IRD *Directrice de thèse*

Anne-Sophie ROBILLIARD

Chargée de Recherche, IRD *Co-encadrante*

Josselin THUILLIEZ

Directeur de Recherche, CNRS- *Examineur*

Université Paris 1 Panthéon Sorbonne

L'Université Paris-Dauphine n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses; ces opinions doivent être considérées comme propres à leurs auteurs.

A

*ma grand-mère Soukeina, mes parents Sidy et Lala,
mon amie Maria-Belén Ojeda Trujillo*

REMERCIEMENTS

Cette thèse est le prolongement d'un stage de Master 2 effectué à DIAL en binôme avec Maria-Belén Ojeda Trujillo dans le cadre du projet DevObs sous la supervision d'Anne-Sophie Robilliard et Charlotte Guénard. Ce stage a été le point de départ et nous a conduits à faire un terrain de recherche passionnant au Sénégal, ce qui a sans nul doute contribué grandement à mon intérêt pour la recherche. Je leur exprime ma profonde gratitude pour avoir guidé mes premiers pas dans la recherche académique et m'avoir fait connaître l'observatoire de population de Niakhar. Elles m'ont également initié au traitement de données démographiques longitudinales qui sont très complexes et pas faciles à prendre en mains. Les deux premiers chapitres de cette thèse portent sur la zone du suivi démographique de Niakhar et ce stage a été d'un grand apport pour la formulation de mes questions de recherche et m'a permis de mieux appréhender les problématiques de développement dans cette zone rurale.

Sandrine Mesplé-Somps qui avait obtenu un financement pour une thèse sur la question des dynamiques de peuplement en Afrique sub-saharienne a accepté de superviser ma thèse avec Anne-Sophie Robilliard. L'enthousiasme, la rigueur scientifique et le sens de la pédagogie dont elles ont fait preuve ont contribué de manière significative à la réalisation de ce travail de recherche. J'ai beaucoup appris à leur côté tant sur le plan du raisonnement scientifique que des techniques quantitatives d'évaluation d'impact. J'apprécie leurs qualités humaines, notamment durant les moments difficiles de cette thèse marqués par deux confinements en 2020, leur soutien a été inestimable. Qu'elles trouvent à travers ces mots, l'expression de toute ma reconnaissance. Cette thèse a bénéficié d'un financement de l'Institut de Recherche pour le Développement (IRD) à travers la bourse ARTS, je leur adresse mes remerciements pour ce financement.

Je tiens à remercier Isabelle Chort et Karine Marazyan d'avoir accepté d'être les rapportrices de ma thèse et d'y consacrer leur temps précieux, je vous en suis très reconnaissant. Les dis-

cussions lors de ma pré-soutenance ont été enrichissantes et vos remarques et suggestions ont permis d'améliorer ce travail. Je remercie également Philippe De Vreyer et Josselin Thuilliez d'avoir accepté de participer à l'évaluation de ce travail. Mes remerciements vont aussi à Elise Huillery, membre de mon comité de suivi de thèse avec Karine Marazyan. Nos échanges dans le fonds sur les différents drafts que j'ai eus à leur présenter durant les comités, leurs encouragements, et leur disponibilité m'ont permis d'approfondir mes analyses et de mener à bien ce travail. Je leur témoigne ma reconnaissance.

J'exprime ma gratitude à Valérie Delaunay, co-auteure sur le premier chapitre de ma thèse pour les discussions très enrichissantes que nous avons eues durant mes différents séjours à l'IRD au Sénégal et ses commentaires éclairés sur le premier draft de ce chapitre. Elle a aussi contribué à ma meilleure compréhension de la structure des données démographiques longitudinales de Niakhar en m'expliquant des astuces et techniques pour mieux explorer ces données. En me donnant son bureau, elle m'a permis de finaliser la rédaction de cette thèse dans de très bonnes conditions. Je lui adresse toute ma reconnaissance, de même que la représentante de l'IRD au Sénégal, Isabelle Henry. Je n'oublie pas Richard Lalou et Laurence Fleury pour leur disponibilité et leurs conseils bienveillants. Je remercie Abdoulaye Diop, un collègue devenu un ami pour ses encouragements constants. Un grand merci à toute l'équipe de l'IRD à Dakar et à Niakhar pour leur accueil chaleureux. J'exprime ma gratitude au service de l'Inspection d'Enseignement et de la Formation (IEF) et au Service Départemental pour le Développement Rural (SDDR) de la région de Fatick du Sénégal pour avoir mis à notre disposition des données sur les infrastructures scolaires et des données pluviométriques.

Ma gratitude va également à toute l'équipe de DIAL pour ces très belles années passées aux côtés de chercheurs inspirants et disponibles, et de mes collègues doctorants très sympathiques. J'ai énormément appris durant les séminaires, workshops, et les discussions informelles sur mes travaux de recherche. Un grand merci pour votre accueil chaleureux, vos qualités relationnelles, votre disponibilité et pour m'avoir permis de travailler dans de très bonnes conditions. J'ai eu un réel plaisir à partager un bureau avec Marin Ferry et j'apprécie sa bonne humeur, sa disponibilité, ses encouragements et surtout son sens de l'humour. Je n'oublie pas les matchs de football organisés par Thomas Calvo et Antoine Boucher, des moments de plein épanouissement. Je remercie Gianluca Orefice, Joachim Jarreau, Jean-Noël Senne, Kenneth Hounghbedji, Thomas Thivillon pour les discussions intéressantes et leurs commentaires éclairés sur mes travaux. Merci à Danielle Delmas, Anne Legendre, Loïc Le Pezennec pour leur gentillesse et leur bonne humeur. J'ai eu une grande joie de voir mes amis Yvan Assany, Komlavi II Adjegan, et Cécile Mouchel avec qui j'ai passé deux très belles années de Master EPOLPRO à l'IEDES me rejoindre

à DIAL en tant que doctorants.

Je tiens à remercier Murray Leibbrandt et Anda David pour leur disponibilité et les discussions enrichissantes que nous avons eues dans le cadre de mon troisième chapitre de thèse qui porte sur l'Afrique du Sud. Je devais partir à l'université de Cape Town pour un séjour de recherche mais ce visiting n'a finalement pas eu lieu en raison du covid.

Un grand merci à mon amie Maria-Belén Ojeda Trujillo avec qui j'ai effectué une mission de terrain de quatre mois au Sénégal pour la collecte de données sur les interventions de développement à Niakhar dans le cadre du projet DevObs. Une personne aux qualités exceptionnelles, j'ai eu une joie ineffable à travailler avec elle durant notre stage. Je lui remercie pour son amitié et ses encouragements tout au long de cette thèse. Merci à mes amies Clara, Fatima, Julie, Rawane, et Sveva pour leur gentillesse et leurs encouragements. Je remercie aussi Akim pour ses conseils bienveillants durant cette thèse.

Enfin, je remercie ma grand-mère pour sa générosité légendaire. Je remercie également mes parents et ma famille qui ont tant investi sur mon éducation. Je leur exprime ma reconnaissance infinie. Votre affection, vos encouragements durant les moments de doute, et votre soutien inestimable ont grandement contribué à la réalisation de ce travail. Une mention spéciale à ma mère, merci infiniment.

Résumé

L'Afrique sub-saharienne fait partie des régions du monde qui posent les plus grands défis présents et à venir de peuplement et de développement du fait tant des dynamiques démographiques observées que des conditions climatiques des plus difficiles. En 2050, la population africaine est estimée à plus de deux milliards, ce qui représentera un quart de la population mondiale. Les travaux empiriques de cette thèse analysent les interactions entre les dynamiques de population, les infrastructures et les migrations internes dans deux pays d'Afrique sub-saharienne en mobilisant des données démographiques longitudinales et de panel originales et relativement rares dans le contexte africain. Le chapitre 1 examine l'effet causal d'une forte augmentation de la construction d'écoles entre 1998-2005 sur la fécondité, le mariage précoce et la santé reproductive dans la zone de l'observatoire de population de Niakhar située en milieu rural sénégalais. Les données du suivi-démographique ont été appariées aux données géo-référencées et datées des infrastructures scolaires nous permettant d'identifier les femmes exposées et non exposées à l'expansion de l'offre éducative. Les résultats montrent que les femmes exposées ont plus de chance d'achever l'école primaire et ont moins d'enfants avant l'âge de 25 ans que les femmes non exposées. Le retard de l'âge au mariage et un changement de comportement sanitaire lié au lieu d'accouchement semblent être les mécanismes à l'œuvre. Accoucher moins à la maison réduit le risque de mortalité infantile, ce qui évite aux mères de s'engager dans une stratégie de fécondité de rattrapage. Cependant, l'exposition aux écoles ne semble pas avoir un effet sur les comportements sanitaires liés aux visites prénatales. Le chapitre 2 étudie l'impact des chocs pluviométriques sur la mortalité infantile en milieu rural sénégalais sur trois décennies 1985-2016 en fusionnant des données pluviométriques géo-localisées et les données du suivi-démographique de Niakhar. Une déviation positive des précipitations par rapport à la moyenne historique augmente le taux de mortalité post-néonatale au cours de la période 1985-2016 avec un effet fort en saison des pluies suggérant un effet maladie-environnement. En analysant l'hétérogénéité de l'impact des pluies selon le sexe, nous trouvons un effet de genre frappant suggérant que seules les filles sont affectées par les précipitations. Par ailleurs, l'ampleur de l'effet des chocs pluviométriques est plus important dans les villages disposant de moins d'infrastructures, ce qui souligne l'importance des infrastructures dans l'atténuation des effets néfastes des chocs climatiques sur la mortalité. Au cours de la dernière décennie, caractérisée par une augmentation des précipitations avec une forte variabilité interannuelle par rapport aux deux premières décennies, l'effet revenu-nutrition domine l'effet maladie-environnement, ce qui suggère que la distribution massive de moustiquaires imprégnées a permis de réduire la mortalité.

ité liée au paludisme. Le troisième et dernier chapitre examine l'effet des migrations internes en Afrique du Sud sur la scolarisation des enfants. En utilisant les données de panel d'individus de National Income Dynamics Study (NIDS) de 2008 à 2017, je compare la progression scolaire et le niveau d'éducation entre les enfants migrants et non-migrants. Les résultats suggèrent que les enfants migrants, notamment ceux qui migrent en même temps avec leur mère ont une meilleure progression scolaire que les enfants non migrants.

Mots clés : Infrastructures, Dynamiques de population, Migrations internes, Education, Changement climatique, Afrique sub-Saharienne.

Abstract

Sub-Saharan Africa is one of the regions of the world that faces the greatest current and future challenges in population and development due to the demographic dynamics observed and the most difficult climatic conditions. In 2050, the African population is estimated to be more than two billion and will represent a quarter of the world population. The empirical works in this thesis analyze the interactions between population dynamics, infrastructures, and internal migrations in two sub-Saharan African countries by mobilizing original longitudinal demographic data and panel data relatively rare in an African context. Chapter 1 examines the causal effect of a sharp increase in school construction between 1998 and 2005 on fertility, early marriage, and reproductive health behavior in the Niakhar Health and Demographic Surveillance System (HDSS) in rural Senegal. The Niakhar HDSS data were matched with geo-referenced and dated school data, allowing us to identify women exposed and not exposed to the expansion of school supply. The results show that exposed women are more likely to complete primary school and have fewer children before the age 25 than unexposed women. A delay in the age at marriage and health behavior change related to childbirth location appear to be the mechanisms at work. Delivering less at home lowers the risk of infant mortality which prevents mothers to engage in a catch-up fertility strategy. However, exposure to schools does not seem to affect health behavior related to prenatal visits. Chapter 2 investigates the impact of rainfall shocks on infant mortality in rural Senegal over three decades 1985-2016 by merging geo-referenced rainfall data and the Niakhar HDSS data. A positive deviation of rainfall from the historical mean increases the post-neonatal mortality rate during the period 1985-2016 with a strong effect in the rainy season suggesting a disease-environment effect. Analyzing the heterogeneity of the impact of rainfall by gender, we find a striking gender effect suggesting that only girls are affected by rainfall. Furthermore, the magnitude of the effect of rainfall shocks is larger in villages with fewer infrastructures stressing that infrastructures matter in the mitigation of adverse effects of climate shocks on mortality. During the last decade characterized by increased precipitations with high inter-annual variability relative to the first two decades, the income-nutrition effect dominates the disease-environment effect suggesting that the massive distribution of Insecticide-Treated mosquito Nets performed well to reduce malaria-related mortality. The third and last chapter examines the effect of internal migrations in South Africa on schooling. Using the National Income Dynamics Study (NIDS) individual panel data 2008-2017, I compare the schooling progression and educational attainment between migrant and non-migrant children. The results suggest that migrant children, especially those who migrate at the same time with their mothers have

better schooling progression than non-migrant children.

Keywords : Infrastructures, Population dynamics, Internal migrations, Education, Climate change, Sub-Saharan Africa.

Contents

0.1	Introduction générale	1
1	Impact of school construction on fertility, early marriage, and reproductive health behavior: Evidence from rural Senegal	7
1.1	Introduction	10
1.2	Background	13
1.3	Data and Descriptive Statistics	15
1.3.1	Demographic data	15
1.3.2	School data	15
1.3.3	Descriptive Statistics	16
1.3.4	Determinants of School Construction	17
1.4	Identification Strategy	19
1.4.1	Difference-in-Differences	19
1.4.2	Threats to Identification	21
1.5	Results	23
1.5.1	Impact on schooling	23
1.5.2	Impact on early fertility, child marriage, and health behavior	26
1.6	Channels and Discussion	28
1.7	Conclusion	30
1.8	Appendix	39
2	Rainfall and Infant Mortality in Rural Senegal	45
2.1	Introduction	50
2.2	How do rainfall fluctuations affect health outcomes and infant mortality ?	52
2.3	Background	57
2.4	Data and Measurement of key variables	59
2.4.1	Mortality outcomes	59

2.4.2	Rainfall data	60
2.4.3	Other Data	61
2.4.4	Descriptive statistics	62
2.5	Identification strategy	63
2.6	Results	65
2.6.1	Main results	65
2.6.2	Heterogeneity	66
2.6.3	Robustness checks	69
2.7	Conclusion	70
3	Children’s migration and Schooling in South Africa	83
3.1	Introduction	90
3.2	Data and descriptive statistics	93
3.2.1	Data	93
3.2.2	Sample	94
3.2.3	Measurement of educational outcomes	94
3.2.4	Migration measurement and spatial inequalities in school quality	95
3.2.5	Baseline characteristics and determinants of children’s migration	96
3.3	Identification strategy	97
3.3.1	Fixed-Effects estimation – within household comparison	97
3.3.2	Cross-section estimation with mother fixed-effects	99
3.4	Results	100
3.4.1	Main results	100
3.4.2	Heterogeneity by gender and type of move	101
3.4.3	Children’s move and maternal migration	103
3.5	Conclusion	104
A6	School inequalities in South Africa	118

List of Figures

1.1	Number of Primary school in Niakhar HDSS 1984-2017	31
1.2	Evolution of the Total Fertility Rate (TFR) in Niakhar HDSS and Rural Senegal 1984-2017	31
1.3	Evolution of the age at first marriage in Niakhar HDSS 1984-2017	32
1.4	Percentage of mothers aged 15-49 years old attended Prenatal visits in Niakhar HDSS 1994-2017	32
1.5	Childbirth location of mothers aged 15-49 years old in Niakhar HDSS 1984-2017 .	33
1.6	Effect of the school construction program on the probability to complete primary school by cohort of birth	33
A1	Percentage of women who attended primary school by year of birth in Niakhar HDSS	39
A2	Fertility rate by cohort in Niakhar HDSS 1984-2017	40
A3	Population aged 6-11 years old in treated and control villages	41
A4	Primary school attendance rate in 1997 (before the sharp increase in school con- struction in Niakhar HDSS)	41
A5	Child mortality rate in the Niakhar area 1990-2016	42
A6	Internal mobility rate within the Niakhar area 1998-2005	42
2.1	Location of Niakhar area and Rain station.	72
2.2	Timing of outcomes	72
2.3	Evolution of Post-Neonatal and Child Mortality rates in the Niakhar HDSS 1985- 2016.	72
2.4	DHS data compared to the Niakhar HDSS data.	73
2.5	Yearly precipitation (in mm) in the 2 communes of Niakhar area.	73
2.6	Trend of mortality outcomes and rainfall over time 1985-2016 in the Niakhar area.	74
2.7	Out-Migration rate of children under 5 per year between 1985-2016.	74
3.1	Cross-section sample	105
3.2	Enrolment rate of black individuals aged 7-20 in each wave from 2008 to 2017 . .	105
3.3	School gap of black children 2008-2017	106

3.4	Age at Grade 12 achievement	106
3.5	Household residency status of black children's parents 2008-2017	107
3.6	Proportion of schools in quintile Q5 at the district level in 2017	107
3.7	Proportion of schools in quintile Q1 at the district level in 2017	108
3.8	Type of move and School progression	108

List of Tables

Chapter 1

Table 1.1	Descriptive Statistics	34
Table 1.2	Probability to be located less than 2km from the school	34
Table 1.3	Effect of the school construction program on the probability to complete primary school	35
Table 1.4	Effect of the school construction program on the probability of Out-Migration	35
Table 1.5	Effect of the school construction program on early Fertility, child Marriage, and Health behavioral outcomes	36
Table 1.6	Robustness check - excluding internal migrants	37
Table 1.7	Exposure to school construction program and Meningitis outbreak 1998-1999	38
Table A1	Effect of the school construction program on the probability to complete primary school - Placebo test	43
Table A2	Effect of the school construction program on early Fertility, child Marriage, and Behavioral outcomes - Placebo test	43
Table A3	Heterogeneous effects by distance to Health posts	44
Table A4	Heterogeneous effects by distance to Secondary schools	44

Chapter 2

Table 2.1	Descriptive statistics	75
Table 2.2	Impact of rainfall on Post-Neonatal and Child Mortality	75
Table 2.3	Impact of rainfall on Post-Neonatal Mortality by Cause of death	75
Table 2.4	Impact of rainfall on Child Mortality by Cause of death	76
Table 2.5	Impact of rainfall on Post-Neonatal and Child Mortality by Gender	76
Table 2.6	Descriptive statistics on post-neonatal and child mortality by cause of death and gender	77
Table 2.7	Heterogeneous effects of rainfall on Post-Neonatal and Child Mortality by distance to Health facilities	77
Table 2.8	Heterogeneous effects of rainfall on Post-Neonatal Mortality by period	78

Table 2.9	Heterogeneous effects of rainfall on Child Mortality by period	79
Table 2.10	Correlation between Rainfall and Crop yields in the department of Fatick 1986-2016	79
Table 2.11	Impact of rainfall on Post-Neonatal and Child Mortality with sample restric- tion: excluding the 3 biggest villages, 1985-2016	80
Table 2.12	Impact of rainfall on Post-Neonatal and Child Mortality with sample restric- tion: excluding 13 villages close to flood zones, 1985-2016	81
Table 2.13	Impact of rainfall on Post-Neonatal and Child Mortality using the level of rainfall, 1985-2016	82

Chapter 3

Table 3.1	Number of black children observed in each wave (2008-2010-2012-2014-2017)	109
Table 3.2	Migration status of black children and their mothers 2008-2017	109
Table 3.3	Type of migration of black children 2008-2017	109
Table 3.4	Descriptive Statistics	109
Table 3.5	Baseline characteristics in 2008 between migrant and non-migrant children .	109
Table 3.6	Determinants of children's migration	110
Table 3.7	Children's migration and School progression	110
Table 3.8	Migration and Grade 12 achievement	111
Table 3.9	Children's migration and School progression - Heterogeneity by gender	111
Table 3.10	Children's migration and School progression - Heterogeneity by gender and type of move	112
Table 3.11	Migration and Grade 12 achievement - Heterogeneity by gender	113
Table 3.12	Children's move and maternal migration	114
Table 3.13	Children's move and maternal migration (Cross-section sample)	115
Table A1	Differences between treated and control groups (Cross-section sample)	116
Table A2	Migration and School progression - Controlling for time-varying shock	116
Table A3	Children's migration and School progression by gender - Controlling for time- varying shock	117
Table A4	Children's migration and School progression by gender and type of move - Controlling for time-varying shock	117
Table A5	Children's move and maternal migration - Controlling for time-varying shock	118

0.1 Introduction générale

La population africaine est estimée atteindre plus de deux milliards en 2050, ce qui représentera un quart de la population mondiale. L'Afrique sub-saharienne contribuera à plus de la moitié de l'augmentation de la population dans le monde dans les trente prochaines années du fait de sa croissance démographique relativement rapide. Cette dynamique démographique observée dans cette région du monde retarde la transition démographique et pose des défis considérables présents et à venir de développement exacerbés par des conditions climatiques des plus difficiles.

Entre 2000 et 2015, des financements s'élevant entre 23 à 30 milliards de dollars étaient prévus pour la construction d'infrastructures scolaires en Afrique sub-saharienne pour la réalisation de l'Objectif du Millénaire pour le Développement (OMD) relatif à l'universalisation de l'éducation primaire (Theunynck 2009). Deux millions de nouvelles salles de classes étaient nécessaires pour atteindre cet objectif, soit 200 000 par an. Les pays de cette région ont alloué plus de 4% de leur Produit Intérieur Brut (PIB) à l'éducation par an en moyenne sur cette période (Majgaard and Mingat 2012). Des progrès ont été faits ces dernières décennies, le taux d'achèvement de l'école primaire est passé de 54% en 2000 à 70% en 2020 dans les pays de la zone sahélienne africaine¹. Cependant, malgré les ressources financières importantes investies dans ces pays pour améliorer l'accès et la qualité de l'éducation, notamment pour les filles, le taux de fécondité précoce y est le plus élevé au monde avec plus de 115 naissances pour 1000 femmes âgées entre 15 et 19 ans en 2018. Or, la scolarisation des filles peut augmenter leur coût d'opportunité d'avoir un enfant et les inciter à réduire leur fécondité (Becker 1992; Schultz 1994; Schultz 2002). L'éducation des filles peut également retarder l'âge au mariage et faire baisser la fécondité précoce. En 2021, 13,3 millions d'enfants, soit 10% du total des naissances dans le monde, sont nés de mères âgées de moins de 20 ans, et la majorité de ces mères se trouve en Afrique sub-saharienne (UNDESA 2022). Cette situation suscite des interrogations sur l'impact des infrastructures scolaires sur la fécondité et le mariage précoce, et la santé reproductive des femmes dans cette région qui fait partie des régions les plus pauvres au monde. Par ailleurs, cette question est d'un grand intérêt pour les décideurs publics qui ont besoin de mieux appréhender dans quelle mesure l'éducation des femmes affecte leur santé reproductive d'une part, et d'évaluer les progrès réalisés en vue de l'atteinte des Objectifs de Développement Durable (ODD) d'ici 2030 d'autre part. Ils ont également besoin de comprendre la dynamique démographique pour une meilleure planification des interventions de développement dans un contexte marqué par une vulnérabilité accrue et sans précédent de l'Afrique sub-saharienne

¹Source: <https://donnees.banquemondiale.org/indicateur/SE.PRM.CMPT.ZS?end=2020&locations=ZG&start=1970&view=chart>

face au changement climatique alors qu'elle y contribue le moins. Entre 1995 et 2015, plus de 40% des catastrophes naturelles liées au climat dans le monde ont été observées dans les pays à revenu faible et intermédiaire occasionnant ainsi plus de 358 000 morts (UNISDR 2015). En outre, un rapport de la Banque mondiale souligne que le changement climatique pourrait conduire 86 millions d'Africains d'origine sub-saharienne à migrer à l'intérieur de leur pays d'ici 2050, et faire basculer plus de 132 millions de personnes dans le monde dans la pauvreté d'ici 2030 (Rigaud et al. 2018).

Cette thèse analyse empiriquement les interactions entre les dynamiques de population, les constructions d'écoles, l'éducation et les migrations internes en Afrique sub-saharienne. Elle mobilise des données démographiques longitudinales et de panel originales et relativement rares dans le contexte africain. Les deux premiers chapitres présentent des travaux portant sur le Sénégal et traitent de l'impact des infrastructures scolaires et des chocs climatiques sur la fécondité et la mortalité infantile dans la zone de Niakhar respectivement, tandis que le troisième traite de l'effet des migrations internes sur l'éducation des enfants et est appliqué à l'Afrique du Sud. La thèse apporte un éclairage nouveau sur la compréhension des dynamiques démographiques tout en mettant en lumière de nouveaux mécanismes. Dans la littérature économique, les travaux qui traitent de la question de l'impact des infrastructures scolaires et de l'éducation sur la fécondité (Osili and Long 2008; Keats 2018; Adu Boahen and Yamauchi 2018; Zenebe Gebre 2020) ou des effets des chocs climatiques sur la mortalité infantile en Afrique sub-saharienne (Kudamatsu, Persson, and Strömberg 2012; Flatø and Kotsadam 2014; Rabassa, Skoufias, and Jacoby 2014; Hyland and Russ 2019) s'appuient essentiellement sur les données Demographic and Health Surveys (DHS). En l'absence d'un système d'Etat Civil complet et performant dans la plupart de ces pays, les données DHS semblent être la source privilégiée des chercheurs pour aborder ces thématiques. Cependant, les enquêtes rétrospectives DHS font l'objet de biais de mémoire occasionnant des omissions de naissances et décès, notamment en milieu rural, de même que des biais d'échantillonnage (Romero Prieto, Verhulst, and Guillot 2021). Le risque du biais de mémoire est d'autant plus élevé que la période séparant la date de l'évènement démographique et l'année de l'enquête est importante. En effet, le biais de sous-estimation du taux de fécondité peut aller jusqu'à 24% en Afrique sub-saharienne, soit 1,5 enfants par femme (Pullum et al. 2013), ce qui pourrait remettre en question la crédibilité des indicateurs démographiques. Par ailleurs, dans le contexte du Sénégal, seules 50% des naissances en milieu rural étaient enregistrées dans le système d'Etat Civil en 2012 (ANSD 2012). Dans une telle situation, les données démographiques longitudinales de l'observatoire de population de Niakhar, situé en milieu rural sénégalais, offrent une opportunité unique pour traiter des

problématiques de développement à une échelle géographique fine. Dans cette zone rurale, des enquêtes démographiques sont réalisées au moins deux fois par an depuis 1983 pour collecter des données sur les événements démographiques tels que les naissances, les décès, les migrations, de même que les mariages, les visites prénatales, les lieux d'accouchement, les causes de décès, permettant ainsi d'avoir des informations d'une précision rare sur une très longue période (De-launay et al. 2013). La fréquence élevée des enquêtes de suivi-démographiques réduit le risque de biais de mémoire et assure une datation précise des événements démographiques par rapport aux données rétrospectives DHS. Les deux premiers chapitres de cette thèse s'appuient sur cette base de données démographiques longitudinales. Concernant l'Afrique du Sud, les données de panel d'individus de National Income Dynamics Study (NIDS) de 2008 à 2017 permettent d'analyser la relation entre les mouvements des enfants avec leur mère à l'intérieur du pays et leur trajectoire scolaire sur une période de dix ans et de prendre en compte l'hétérogénéité individuelle inobservée dans la partie de l'identification.

Le premier chapitre examine l'effet causal d'une forte augmentation de la construction d'écoles entre 1998-2005 sur la fécondité, le mariage précoce et la santé reproductive. Il existe peu de travaux en économie qui examinent l'effet de la construction d'infrastructures scolaires sur la santé reproductive des femmes, en particulier les visites prénatales et les lieux d'accouchements. Cela tient au fait que de telles données sont rarement disponibles en Afrique sub-saharienne, notamment en milieu rural. En plus de réduire la fécondité précoce avant l'âge de 25 ans, la construction d'écoles conduit à des changements de comportements sanitaires. Les résultats mettent en lumière un nouveau mécanisme jusque-là pas montré dans la littérature indiquant que le fait d'accoucher moins dans les domiciles réduit le risque de mortalité infantile, ce qui évite aux femmes de s'engager dans une stratégie de fécondité de rattrapage.

Le deuxième chapitre étudie l'impact des chocs pluviométriques sur la mortalité infantile à Niakhar sur trois décennies 1985-2016 en mobilisant des données sur les causes de décès pour mieux comprendre les mécanismes à l'œuvre. Le degré de précision de la datation des événements démographiques grâce à la fréquence élevée des enquêtes de suivi démographique permet de distinguer les effets de maladie-environnement et revenu-nutrition. Il montre qu'une déviation positive des précipitations par rapport à la moyenne historique augmente le taux de mortalité post-néonatale au cours de la période 1985-2016 avec un effet fort en saison des pluies suggérant un effet maladie-environnement. En analysant l'hétérogénéité de l'impact des pluies selon le sexe, nous trouvons un effet de genre frappant suggérant que seules les filles sont affectées par les précipitations. Par ailleurs, l'ampleur de l'effet des chocs pluviométriques est plus importante dans les villages disposant de moins d'infrastructures sanitaires, hydrauliques, élec-

triques, ce qui souligne l'importance des infrastructures dans l'atténuation des effets néfastes des chocs climatiques sur la mortalité. Au cours de la dernière décennie, caractérisée par une augmentation des précipitations avec une forte variabilité interannuelle par rapport aux deux premières décennies, l'effet revenu-nutrition domine l'effet maladie-environnement, ce qui suggère que la distribution massive de moustiquaires imprégnées a permis de réduire la mortalité liée au paludisme.

Le troisième et dernier chapitre examine l'effet des migrations internes en Afrique du Sud sur la scolarisation des enfants. Jusqu'à présent, les travaux dans la littérature économique ont prioritairement étudié l'effet de la migration de travail des parents, en particulier à l'international, sur l'éducation des enfants dans le ménage d'origine (Amuedo-Dorantes and Pozo 2010; Antman 2011; Antman 2012; Meng and Yamauchi 2017). De ce fait, il y a un manque de connaissances sur la trajectoire scolaire des enfants qui accompagnent leur mère qui migre à l'intérieur du pays pour des opportunités de travail. Or, certaines sociétés d'Afrique sub-saharienne sont caractérisées par une mobilité géographique interne forte des enfants, comme c'est le cas en Afrique du Sud. Ce chapitre tente d'explorer cette question qui est très peu traitée dans la littérature en utilisant les données de panel d'individus de National Income Dynamics Study (NIDS) de 2008 à 2017. Je compare la progression scolaire et le niveau d'éducation entre les enfants migrants et non-migrants dans un contexte où il existe de fortes inégalités spatiales sur la qualité des infrastructures scolaires. Les résultats suggèrent que les enfants migrants, notamment ceux qui migrent en même temps avec leur mère ont une meilleure progression scolaire que les enfants non migrants.

REFERENCES

- Adu Boahen, Emmanuel and Chikako Yamauchi (2018). “The effect of female education on adolescent fertility and early marriage: evidence from free compulsory universal basic education in Ghana”. *Journal of African Economies* 27.2, pp. 227–248.
- Amuedo-Dorantes, Catalina and Susan Pozo (2010). “Accounting for remittance and migration effects on children’s schooling”. *World development* 38.12, pp. 1747–1759.
- ANSD (2012). *Enquête Démographique et de Santé à Indicateurs Multiples au Sénégal (EDS-MICS) 2010-2011*.
- Antman, Francisca M (2011). “The intergenerational effects of paternal migration on schooling and work: What can we learn from children’s time allocations?” *Journal of Development Economics* 96.2, pp. 200–208.
- (2012). “Gender, educational attainment, and the impact of parental migration on children left behind”. *Journal of Population Economics* 25.4, pp. 1187–1214.
- Becker, Gary S (1992). “Fertility and the economy”. *Journal of Population Economics* 5.3, pp. 185–201.
- Delaunay, Valerie et al. (2013). “Profile: the Niakhar health and demographic surveillance system”. *International Journal of Epidemiology* 42.4, pp. 1002–1011.
- Flatø, Martin and Andreas Kotsadam (2014). *Droughts and gender bias in infant mortality in Sub-Saharan Africa*. Tech. rep. Memorandum.
- Hyland, Marie and Jason Russ (2019). “Water as destiny—The long-term impacts of drought in sub-Saharan Africa”. *World Development* 115, pp. 30–45.
- Keats, Anthony (2018). “Women’s schooling, fertility, and child health outcomes: Evidence from Uganda’s free primary education program”. *Journal of Development Economics* 135, pp. 142–159.

- Kudamatsu, Masayuki, Torsten Persson, and David Strömberg (2012). “Weather and infant mortality in Africa”. *CEPR Discussion Paper No. DP9222*.
- Majgaard, Kirsten and Alain Mingat (2012). *Education in Sub-Saharan Africa: A Comparative Analysis*. Washington, D.C.: World Bank. DOI: 10.1596/978-0-8213-8889-1.
- Meng, Xin and Chikako Yamauchi (2017). “Children of migrants: The cumulative impact of parental migration on children’s education and health outcomes in China”. *Demography* 54.5, pp. 1677–1714.
- Osili, Una Okonkwo and Bridget Terry Long (2008). “Does female schooling reduce fertility? Evidence from Nigeria”. *Journal of Development Economics* 87.1, pp. 57–75.
- Pullum, Thomas W et al. (2013). “An assessment of DHS estimates of fertility and under-five mortality”. In: *International Population Conference of the International Union for the Scientific Study of Population (IUSSP), Session 132: Data quality in demographic surveys, August*. Vol. 28.
- Rabassa, Mariano, Emmanuel Skoufias, and Hanan Jacoby (Apr. 2014). “Weather and Child Health in Rural Nigeria”. *Journal of African Economies* 23.4, pp. 464–492. ISSN: 0963-8024. DOI: 10.1093/jae/eju005. eprint: <https://academic.oup.com/jae/article-pdf/23/4/464/2241888/eju005.pdf>.
- Rigaud, Kanta Kumari et al. (2018). *Groundswell : Preparing for Internal Climate Migration*. World Bank, Washington, DC.
- Romero Prieto, Julio, Andrea Verhulst, and Michel Guillot (2021). “Estimating the infant mortality rate from DHS birth histories in the presence of age heaping”. *PloS one* 16.11, e0259304.
- Schultz, T Paul (1994). “Human capital, family planning, and their effects on population growth”. *The American Economic Review* 84.2, pp. 255–260.
- (2002). “Why governments should invest more to educate girls”. *World development* 30.2, pp. 207–225.
- Theunynck, Serge (2009). *School construction strategies for universal primary education in Africa: should communities be empowered to build their schools?* World Bank Publications.
- UNDESA (2022). *World Population Prospects 2022: Summary of Results*. UN DESA/POP/2022.
- UNISDR (2015). *Weather Disasters Report: The human cost of weather related disasters 1995-2015*.
- Zenebe Gebre, Tihtina (2020). “Free Primary Education, Timing of Fertility, and Total Fertility”. *The World Bank Economic Review* 34.3, pp. 730–748.

CHAPTER 1

IMPACT OF SCHOOL CONSTRUCTION ON FERTILITY, EARLY MARRIAGE, AND REPRODUCTIVE HEALTH BEHAVIOR: EVIDENCE FROM RURAL SENEGAL

Abstract¹

Sub-Saharan Africa still displays the highest early fertility rate in the world while considerable funds have been invested to improve both school infrastructure supply and women's education. In this paper, we investigate the causal effect of a sharp increase of school construction during the period 1998-2005 in Niakhar, rural Senegal on demographic and reproductive health behavior outcomes. We use longitudinal and high-quality demographic data combined with geo-referenced school data. Exploiting both spatial and temporal variations in school construction, we use a difference-in-differences identification strategy. We find that school construction between 1998 and 2005 increases the probability to complete primary school by 10.4 percentage points. Exposure to school construction during that period reduces early fertility before age 25 by 39%. Schooling decreases early fertility through delaying the age at marriage and reproductive health behavior related to childbirth location. Delivering in a health facility lowers the risk of infant mortality which could prevent mothers to engage in a catch-up fertility strategy. Schooling does not seem to affect health behavior related to prenatal care. We provide evidence indicating that out-migration and internal mobility, the two main threats for the identification, do not drive our results.

JEL Classification: I12, I25, J12, J13.

Keywords: Education, Fertility, Marriage, Health Behavior, Niakhar HDSS, Sub-Saharan Africa.

¹This chapter is co-authored with Valérie Delaunay.

1.1 Introduction

Sub-Saharan Africa experienced a significant rise in school infrastructures after the “World Conference on Education for All” in Jomtien (Thailand) in 1990 and the “World Education Forum” in Dakar (Senegal) in 2000. The proportion of children completing primary school in this region increased from 46% to 68% between 1970 and 2010 while the share of children who complete secondary school grew from 22% to 40% over the same period (Evans and Mendez Acosta 2021). Despite substantial gains in schooling over the last decades, the region displays the highest early fertility rate in the world, with about 115 births per 1,000 women aged 15-19 in 2018, combined with a low prevalence of prenatal care before delivery². Child marriage still matters as 115 million girls get married before age 18 in 2018 (UNICEF 2018) and one out of three women married by age 18 in 2012 in developing countries (McGavock 2021). Chari et al. (2017) show that early marriage adversely affects child health outcomes, increases complete fertility, and constrains women empowerment. This pattern leads researchers to try to better understand to what extent school infrastructures and more broadly women’s schooling might affect early childbearing, child marriage, and child health outcomes in the developing world.

Causal evidence of schooling on early fertility and child marriage in sub-Saharan exists but the mechanisms at work vary and seem to be context specific. Using Demographic and Health Survey (DHS) data, Osili and Long (2008) investigate the impact of a massive school construction program between 1976 and 1981 in Nigeria on fertility before age 25 years old. The authors estimate a difference-in-differences model by exploiting both intensity and timing of the program and rely on an instrumental variable method to overcome the endogeneity issue of education. Exposure to the program is used as an instrument for education and IV estimates indicate that one additional year of schooling leads to an 11% to 19% reduction in early fertility. However, the authors do not provide evidence on channels through which schooling affects early fertility. Taking advantage of the implementation of the universal primary education reform in 1997 in Uganda, Keats (2018) uses a regression discontinuity method to examine the causal effect of women’s education on fertility and child health outcomes. The author uses DHS data and finds that women who benefited from the reform are less exposed to early childbearing. Postponement of age at marriage and age at first birth, contraceptive use, and increased labor market participation are the main pathways through which schooling affects fertility. Adu Boahen and Yamauchi (2018) exploit the implementation of the free compulsory universal basic education policy in Ghana in 1996 as a natural experiment and perform a regression discontinuity analysis to investigate the

²These figures come from UNICEF Data on early childbearing (www.data.unicef.org/topic/child-health/adolescent-health)

impact of schooling on early childbearing and child marriage. The authors rely on four waves of DHS data and find that the reform reduced early fertility and child marriage among women aged 15-20 years old. They interpret the results as a knowledge effect related to the use of contraceptives rather than an abstinence effect. Using a difference-in-differences method and DHS data, Zenebe Gebre (2020) investigates the impact of free primary education reform in the 1990s in Malawi on early fertility before age 25. She finds strong evidence on contraceptive use, birth spacing, delaying the age of first marriage, and age at first birth as the key pathways explaining lower fertility of women exposed to the reform but no effect on labor market participation. In contrast, Chicoine (2020) investigates the impact of a similar reform in Ethiopia and shows that fertility decline is driven by an increased women labor market participation and fewer desirable family size but no evidence on contraceptive use. Taking advantage of the exogeneity of the policy reform in 1995, the author instruments the years of schooling by exposure to the reform and finds that schooling reduces early childbearing. On the other hand, a number of studies focus on the impact of secondary education on teenage pregnancies and child marriage in sub-Saharan Africa (Duflo, Dupas, and Kremer 2015; Ozier 2018). Overall, results show that attending secondary school delays early fertility and child marriage. However, evidence of the impact of mother's schooling on health behavior related to prenatal care, healthcare at birth, and childbirth location are very scarce in sub-Saharan Africa due to data limitations. Given that prenatal visits and childbirth location seem to play a key role in both maternal and infant mortality in low-income countries, examining the effect of education on health behavior at birth is critical for policymakers. To our knowledge, only Keats (2018) analyzes the causal relationship between women's education and child health outcomes at birth in sub-Saharan Africa. The author shows that one additional year of schooling of the mother leads to a 29 percentage points increase in the probability to get a trained assistant at the birth of the first child in Uganda. Looking at childbirth location, he finds no effect of mother's schooling on the probability of a woman delivers her first child at home rather than in health facilities.

This paper investigates the causal effect of a sharp increase in primary school construction between 1998 and 2005 on early fertility, child marriage, and reproductive health behavior related to prenatal care and childbirth location in the Niakhar Health and Demographic Surveillance System (HDSS). Located in rural Senegal, the Niakhar HDSS is one of the oldest HDSS in sub-Saharan Africa. Demographic events such as births, deaths, migrations, marriages, and health behaviors related to childbirth location, prenatal visit, vaccination have been recorded at least twice a year in the thirty villages of the study area over the last three decades, resulting in a longitudinal and high-quality demographic data base which is relatively rare in an African

context. In addition to demographic data, we use geo-referenced school infrastructure data and administrative data collected during a research field in the Niakhar area in 2018. This article contributes in several ways to the existing literature analyzing the causal relationship between primary schooling and demographic outcomes as well as reproductive health behavior. We provide new evidence of the impact of women's schooling on health behavior related to prenatal care and childbirth location while most of the studies in the literature focus on fertility and marital outcomes. Furthermore, taking advantage of the longitudinal demographic data, we examine the long-term trend over three decades of demographic and health outcomes to shed light on the pathways at work. Last, we provide causal evidence of the impact of school infrastructures on education, demographic, and reproductive health behavior outcomes in rural Senegal where, to the best of our knowledge, there is very little empirical work on this research question.

Exploiting both spatial and temporal variations of school construction between 1998 and 2005 in the Niakhar area, we use a difference-in-differences approach to identify the causal effect of school infrastructures on several outcomes of interest. An important threat for the identification strategy is out-migration given that 54% of women born in the area move out. Women exposed to the sharp increase in school construction during the period 1998-2005 might be more educated and out-migrate from their home village to the urban area looking for job opportunities. Since individuals who left the Niakhar area more than six months are not tracked, the outcomes of interest related to early fertility, child marriage, and reproductive health behavior are censored for migrants. The outcomes are only observed for non migrant women leading to a sample selection bias. To analyze to what extent out-migration could constraint the identification strategy, we estimate the effect of the treatment on the educational outcome for both non-migrant and migrant women and the probability of out-migration. We provide evidence indicating that out-migration is orthogonal to the school construction during the period 1998-2005 in the study area. As a result, the attrition of migrants should not bias our estimates even. Furthermore, the internal mobility rate within the area is very low, about 2%, and results remain relatively stable when we exclude internal movers. In addition, we perform placebo tests to make sure that our results are not driven by pre-existing differential trends. The proximity to health facilities could be a potential confounding factor as women might benefit from awareness campaigns on family planning and more broadly on reproductive health. When we interact the treatment variable with the distance to health infrastructures, we show that this pathway does not seem to operate.

Results indicate that school construction between 1998-2005 in the Niakhar area raised the probability to complete primary school by 10.4 percentage points for women aged 3-9 years old in 1998 compared to the older cohort aged 13-18 years old in 1998 that were not exposed. Ex-

posure to the sharp increase in school infrastructures between 1998-2005 reduces the number of births before age 25 by 0.52 corresponding to a 39% decrease in early childbearing relative to the mean in the control group. Delayed age at marriage is one of the main channels through which educational attainment appears to affect early fertility. Exposure to schools during that period decreases the probability of marriage before age 18 by 16.2 percentage points. Moreover, women exposed to school in childhood are less likely to deliver their first child at home entailing a lower risk of infant mortality for kids. Lower infant mortality could in turn prevent mothers from engaging in a catch-up fertility strategy. We find no evidence of the impact of exposure to school infrastructures on prenatal care.

The remainder of the paper is organized as follows. Section 2 provides background on education policy in Senegal and the study area. Section 3 presents the data and descriptive statistics and stresses the potential determinant of school construction. Section 4 explains the identification strategy. Section 5 shows the main results and the heterogeneity of the effect by cohorts of birth. Section 6 discusses the mechanisms at work and Section 7 concludes.

1.2 Background

One of the main goal of the “World Conference on Education for All” which took place in Jomtien (Thailand) in 1990 was to reach universal access to school for all children in developing countries by the year 2000. Although the important progress in school supply in developing countries, more than 113 million children have no access to primary education in 2000 (UNESCO 2000). Then, Dakar (Senegal) hosted the World Education forum in 2000 in order to promote a collective commitment in developing countries, especially in sub-Saharan countries to meet the universal access to primary school for all children by 2015.

These two conferences are important milestones in the process of improving both school access and quality of education in Senegal. Between 1994 and 1998, more than 2,000 classrooms in the primary school were built in Senegal, and the government allocated about US \$2 million of the Senegalese’s budget for school construction in 1998 (Theunynck 2009). Furthermore, a new education policy called PDEF “Programme Décennal de l’Education et de la Formation” was designed by policymakers to be implemented from 2000 to 2011. The objective of this program was twofold: improved school access for all children at the national level and education quality by increasing the number of teachers in primary school. Between 2000 and 2005, 10,214 primary classrooms were built in Senegal and more than 2,000 teachers were hired yearly (WorldBank 2006). The gross enrollment rate in the primary school increased significantly from 57.1% in

1990 to 67.2% in 2000 and reached 93.9% in 2011 at the national level (Diagne [2012](#)). According to the PDEF report (Diagne [2012](#)), public education spending represented 5% of Senegal's GDP on average over the period during the implementation of this program, and the yearly budget allocated to education increased from US \$188 million to US \$776 million between 2000 and 2011.

The large investment in education improved school access in Senegal significantly both in urban and rural areas. In this paper, we focus on the Niakhar Health and Demographic Surveillance System (HDSS), one of the oldest HDSS in Africa created in 1962 by the French National Research Institute for Sustainable Development (IRD) (Delaunay et al. [2013](#)). It is located in rural Senegal in the district of Fatick, 135 km East from the capital Dakar. The size of the Niakhar area is about 203 km². The main driving factor in establishing the Niakhar HDSS was the limited availability of reliable demographic data in Senegal after independence due to the lack of vital registration systems, especially in rural areas³. The population of the Niakhar HDSS has been surveyed at least twice a year and demographic events have been recorded in each wave of the survey. The study area was extended from eight villages in 1962 to thirty villages in 1983. Over the last three decades, the population has increased two-fold from 23,200 in 1984 to 48,200 in 2017.

Agriculture is the main source of income in the Fatick district and the sector employs about 90% of the active population (ANSD [2016](#)). However, seasonal migration is an important coping strategy to deal with income fluctuations (Lalou and Delaunay [2015](#)). Concerning human capital, 82% of household heads in the study area reported in 2014 that they never attended primary school. More than 80% of women aged between 15 and 24 years old in 1990s have never attended primary school and only 1.5% of women use contraceptives in the Niakhar area in 1999 (Adjagbo and Delaunay [2018](#)). The Senegalese government implemented several family planning programs during the last decade in order to improve the access of contraceptives in health facilities, especially in rural areas. According to the 2017 Demographic and Health Survey (DHS), 27.3% of women aged between 15 and 49 in the Fatick district use contraceptives (ANSD [2018](#)).

³Although the vital registration system has improved significantly in Senegal during the last two decades, births and deaths are still widely under reported. According to the Demographic and Health Survey, only 50% of births in rural Senegal have been recorded in a vital registration database compared with 78% in urban areas (ANSD [2012](#)).

1.3 Data and Descriptive Statistics

1.3.1 Demographic data

The demographic data comes from the Niakhar HDSS where demographic events such as births, deaths, migrations, marriages, childbirth locations have been recorded at least twice a year in the thirty villages of the study area since 1983. A census was conducted in 1983 in the thirty villages (Delaunay et al. 2013) allowing the construction of the baseline dataset which contains all Niakhar residents with their individual characteristics such as sex, religion, ethnic group, caste, marital status. Then, the baseline dataset is updated at least twice a year thanks to the demographic surveys conducted in the thirty villages each year under the supervision of the IRD research team based in Senegal. All Niakhar residents are followed and this longitudinal dataset allows tracking the demographic dynamics over a very long period of time as well as all demographic events that occurred during their lifetime. The high frequency of demographic surveys in the Niakhar area over a very long time period lowers the recall bias in demographic events during the interview allowing a high degree of precision in dating events compared to Demographic and Health Surveys (DHS). Concerning reproductive health related variables, surveyors have asked questions related to prenatal visits of women who gave birth in the area since 1994. Also, the level of education of the population has been recorded in several waves of the demographic survey.

The demographic dataset contains information on the village and the year of birth of each individual. This information is key to identify the causal impact of school construction program on outcomes of interest. Taking advantage of the very long time span of data, we also examine the long-term trend of the total fertility rate, fertility rate by age group, the age at first marriage, prenatal visits, and childbirth location in the thirty villages of the area over the last three decades.

1.3.2 School data

During a research field from April to July 2018 in Senegal, Maria-Belén Ojeda Trujillo⁴ and Hamidou Diallo collected data on development interventions (health, education, agriculture, infrastructures) through the “DevObs project” under the supervision of Anne-Sophie Robilliard⁵ and Charlotte Guénard⁶. The purpose of the project was to assess the impact of development

⁴Maria-Belén Ojeda Trujillo and Hamidou Diallo were research interns at DIAL-IRD from March to September 2018.

⁵Research Fellow at IRD, UMR LEDA, Team DIAL.

⁶Professor of Economics at IEDES-Université Paris 1 Panthéon-Sorbonne.

interventions on demographic dynamics in the Niakhar HDSS. During the field work, more than 230 individual interviews were conducted with local development actors such as ministries, local and international NGOs, technical and financial partners, heads of educational and health institutions, and farmers associations, allowing the collection of a rich quantitative and qualitative dataset.

The school infrastructure data is provided by the *Inspection d'Enseignement et de la Formation (IEF)* of Fatick and Diofior which is a service of the Senegalese Ministry of Education in charge of the monitoring of administrative data on primary schools. The data contains the number of primary schools, the school name, and the date of school construction in each village of the Niakhar area between 1950 and 2017. Furthermore, we interviewed the inspectors of the IEF to better understand the determinants of school construction in villages. We then geo-referenced all primary schools located in the thirty villages of the Niakhar HDSS. The demographic and school data are merged using the location of the village of birth.

1.3.3 Descriptive Statistics

Figure 1.1 presents the evolution of the number of primary schools in the Niakhar HDSS during the last three decades. The number of primary schools increased from 4 to 30 between 1984 and 2017 with a sharp increase since 1998. Figure A1 in the appendix shows the share of women who attended primary school by year of birth. The youngest cohort who was born after 1985 and aged 12 years old or less in 1998 has the highest primary school attendance rate. The sharp increase in school construction observed since 1998 in the Niakhar area seems to affect the youngest cohort. Although the official primary school age in Senegal is between 6 and 11 years old, children may enter primary school late and stay in beyond the official age.

Figure 1.2 shows the evolution of the Total Fertility Rate (TFR) in the Niakhar HDSS over the last three decades. The TFR decreased from 8.10 children per woman in 1984 to 5.47 children per woman in 2017 which is slightly below the TFR in rural Senegal in 2017 (5.9 children per woman). At the national level, the TFR declined slightly in Senegal from 6.4 to 4.6 children per woman between 1986 and 2017 (ANSD 2018). Analyzing the number of children per woman in different age brackets over time, we find that a significant decline in early childbearing between 15-19 and 20-24 years old. This is consistent with Garenne et al. (2018) who finds that early fertility decline to be the main driver of TFR reduction in the Niakhar area. As shown in Figure A2 in the appendix, the number of children per woman aged between 15 and 19 decreased from 196 children per 1,000 women to 50 children per 1,000 women between 1984 and 2017 corresponding to a 74% decrease over this period. A similar pattern occurred for women

aged 20 to 24 which experienced a decline in the fertility rate from 309 to 168 children per 1,000 women between 1984 and 2017. Instead, the fertility rate of older cohorts remains high and stable throughout the period. Figure 1.3 presents the evolution of the average age at first marriage in the Niakhar area for women and men. Between 1984 and 2017, the age at first marriage of residents⁷ increased from 18 to 21 for women (16%) and from 25 to 28 for men (12%). Looking at the evolution of reproductive health behavior, we show in Figure 1.4 that the percentage of women aged 15-49 years old who did not attend any prenatal visits declines from 24% to 1% between 1994-2017 while those who attended four visits at least increases from 4% to more than 10% during the same period. The Senegalese Ministry of Health recommends women to attend at least four prenatal visits before delivery but it seems that roughly 90% of women aged 15-49 years old in 2017 did not do so in the Niakhar area. On the other hand, Figure 1.5 indicates that the share of women aged 15-49 years old delivering at home and in health facilities decreased from 88% to 44% and increased from 12% to 52% respectively between 1984 and 2017.

Table 2.1 reports summary statistics for women aged 15-49 years old in 2017 and born in the Niakhar area. The level of education of women is low, 43% attended primary school, and 30% completed it. In Table 2.1, descriptive statistics on fertility, marital, and reproductive health behavior outcomes for non-migrant women are reported. Indeed, once individuals move-out more than 6 months, they are not followed anymore by the HDSS (Lalou and Delaunay 2015). The out-migration rate is high as 54% of women born in the Niakhar area left the area and the average age of out-migration is 15 years old. The three main causes of out-migration are family reasons (37%), work (32%), and marriage (18%). The capital Dakar is the main destination of migrants (39%) but women also migrate in other urban areas as well as in neighboring villages located outside the study area. Sample selection issues will be discussed in Section 4.

1.3.4 Determinants of School Construction

Given that the construction of infrastructures is not randomly assigned, it is useful to try and analyze the rationale behind the construction of primary schools in the Niakhar area before discussing the identification strategy. According to the inspectors of IEF, the target of the Ministry of Education is to build a primary school within 2 km of each village. Hence, when a village has no school, a primary school should be built in a buffer zone of 2 km. According to the inspectors, the decision to build a school depends on both the population of children of schooling age 6-11 years old and the population density of villages. Our identification strategy relies on comparing

⁷Residents include individuals who were born in the Niakhar area and those who became residents through immigration in the study area.

villages than benefit from a nearby school - i.e. constructed less than 2 km from the centroid of the village - and those who do not. Using the sharp increase in construction that started in 1998, four groups of villages can be defined:

- 9 villages had no nearby school before 1998 and benefited from a nearby school construction between 1998 and 2005 (group T1).
- 5 villages had a nearby school before 1998 and benefited from a nearby school construction between 1998 and 2005 (group T2).
- 3 villages had no nearby school before 1998 and did not benefit from a nearby school construction between 1998 and 2005 (group C1)
- 13 villages had a nearby school before 1998 and did not benefit from a nearby school construction between 1998 and 2005 (group C2).

Although school construction continues after 2005, we focus on school construction between 1998 and 2005 in the Niakhar area as these years correspond to the cohorts most exposed to the program and for which outcome variables are observed.

Using the Niakhar demographic database, we compute the number of children of schooling age (between 6 and 11 years old) and the density of population at the village level between 1984 and 2005. In Table 1.2, the probability of benefiting from a nearby school is regressed on these two village characteristics. Results show that this probability is positively correlated with both the population of schooling age and the density of the population in columns (1) and (4). The coefficients increase in magnitude and are still statistically significant when accounting for time-invariant observed and unobserved characteristics of villages as shown in columns (2) and (5). However, once we control for year fixed-effect, the coefficients become negative indicating that as time goes by, smaller villages have a higher probability to have a school. This suggests that primary schools are allocated first to the biggest villages and later to smaller ones.

Indeed, Figure A3 in the appendix is consistent with this rationale. The graph shows the evolution of the population of schooling age between 1984 and 2017 in the four groups of villages mentioned above. Villages with a nearby school before 1998 (C2, orange plot and T2, red plot) have a bigger population of schooling age compared to the two other groups of villages. Comparing villages with a nearby school constructed between 1998 and 2005 (T1, blue plot) and those that still had no school in 2005 (C1, green plot), we find that the latter group is composed of much smaller villages.

In addition, we compute the primary school attendance rate in 1997 before the sharp increase in school construction in the Niakhar area among women aged more than 12 years old in 1997. The intuition behind Figure A4 in the appendix is to assess whether there are differences in primary school attendance before 1998 between the four groups of villages. The figure indicates that the five villages with a nearby school before 1998 and that benefited from an additional school construction between 1998 and 2005 had already higher primary school attendance rate (28,37%) in 1997 compared to other villages. Keeping in mind the size of the five villages, we can assume that the motive of the new school construction during the period 1998-2005 in these villages is to reduce a possible congestion effect in schools. On the other hand, villages exposed to a nearby school after 1998 (blue plot) have a lower primary school attendance rate (11.35%) compared to villages exposed to a nearby school before 1998 (red and orange plots). To sum-up, the population of schooling age, the density of population as well as the primary school attendance seem to matter in the decision of school construction in villages. In the identification strategy, we take into account these pre-existing differences between villages.

1.4 Identification Strategy

1.4.1 Difference-in-Differences

Exploiting both spatial and temporal variations in school construction in the Niakhar area, we use a difference-in-differences method as in Duflo (2001) to identify the causal effect of the school construction program on outcomes of interest. This identification strategy is used in the economic literature to assess the causal impact of school construction program and free primary education reform on a wide range of outcomes (Breierova and Duflo 2004; Osili and Long 2008; Lucas and Mbiti 2012; Zenebe Gebre 2020). As shown in Figure 1.1, the Niakhar area experienced a sharp increase in school construction since 1998. We focus on the sharp increase in school construction during the period 1998-2005. Woman's exposure to school construction depends on both village of birth and year of birth. We compare women born between 1989 and 1995 (aged 3-9 in 1998) with women born between 1980 and 1985 (aged 13-18 in 1998). The youngest cohort is expected to be exposed to the sharp increase in school construction during the period 1998-2005. Instead the older cohort is expected to be out of the age of primary schooling and not exposed. The official primary school age in Senegal is between 6 and 11 years old but some children may enter primary school late and stay beyond the official age due to the grade repetition. According to the monitoring schooling data of the Senegalese Ministry of Education, the grade repetition in primary schools averaged 14% between 1998 and 2000 (PASEC 2004).

We estimate the following difference-in-differences model:

$$Y_{ivc} = \alpha_0 + \beta Young_c * T_v + \gamma X_i + \lambda M_{vc} + \theta_v + \delta_c + \epsilon_{ivc} \quad (1.1)$$

where Y_{ivc} is the outcome of interest of woman i , born in village v and belonging to cohort c , α_0 is a constant, $Young_c$ is a dummy variable equal to 1 if the woman was aged 3-9 in 1998, and 0 if woman was aged 13-18 in 1998, T_v is a dummy variable indicating whether the village of birth was located less than 2km to primary school between 1998-2005 (henceforth "exposed to a nearby school after 1998"), X_i refers to individual time-invariant individual characteristics such as religion and caste, M_{vc} are characteristics of the village of birth in 1997 (schooling age population, schooling rate, population density) interacted with year of birth, θ_v is a village of birth fixed-effect, δ_c is a cohort of birth fixed-effect, ϵ_{ivc} is the error term. β is our coefficient of interest, it captures the impact of the school construction program on outcomes of interest under the parallel trend assumption, i.e. assuming that, in the absence of the program, time trends of outcomes of interest would be the same in the treated and control groups.

Duflo (2001) computes the number of schools built per 1,000 children to measure the intensity of school construction program in Indonesia. Here, we rely on the distance between school and village of birth to capture the exposure to school construction between 1998-2005 in the area. Indeed, the distance between villages and schools averaged 2.5km before 1998 while the mean distance is only 1.1km in 2005 corresponding to a 56% decrease.

As shown in the previous section, the population of schooling age, the density of the population as well as the primary school attendance rate of the village are related to the probability of getting a school. M_{vc} controls for the village-specific characteristics before the sharp increase in school construction in the Niakhar area by interacting each of these variables with the year of birth. Taking into account the pre-existing trends allows us to deal with confounders i.e. factors correlated with both school construction and outcomes of interest. The point estimates are unlikely to reflect a continuation of pre-existing differences. Thus, a causal impact of school construction program on outcomes of interest is identified under the assumption that there are no omitted time-varying variables at the village level during the period 1998-2005 that might be correlated with the treatment. This assumption is very likely to hold as no other public policy shocks correlated with school construction were implemented in the Niakhar area during that period. In interviews conducted during the field work, inspectors stressed that education was the priority of the government during that period. Nevertheless, what remains to be discussed as a likely confounder is the improvement of health facilities in the study area, and awareness cam-

paigns on reproductive health. The Niakhar area has four health infrastructures constructed in 1953, 1957, 1983, 2014 and located outside the nine treated villages. The three oldest health facilities were renovated by international NGOs in 2008, 2009, and 2014 to improve the quality of health services. In addition, the Senegalese government implemented family planning and reproductive health programs in health facilities to promote the use of contraceptives. When we account for the distance to health posts, we provide evidence indicating that our results are not driven by this confounder.

In order to test empirically the validity of the parallel trend assumption, we run a placebo test by comparing women born between 1979 and 1984 (aged 14-19 in 1998) with women born between 1973 and 1978 (aged 20-25 in 1998). Both cohorts are not expected to be exposed to the sharp increase in school construction during the period 1998-2005 in the Niakhar area. The point estimates should be close to zero and statistically insignificant if this assumption holds.

The coefficient of interest β in equation (1) captures the overall effect of school construction on all exposed women aged 3-9 years old in 1998. To analyze the heterogeneity of the effect of the program by cohorts of birth, we estimate the following model:

$$Y_{ivc} = \alpha_0 + \sum_{k=3}^{k=12} \beta_k (d_{ik} * T_v + \gamma X_i + \lambda M_{vc} + \theta_v + \delta_c + \epsilon_{ivc}) \quad (1.2)$$

where d_{ik} is a dummy variable indicating whether individual i was age k in 1998. β_k captures the effect of the sharp increase in school construction between 1998 and 2005 on outcomes of each cohort of women aged from age 3 to 12 in 1998. We expect that the youngest cohort will be more impacted by the school construction during the period 1998-2005 compared to the older cohort.

1.4.2 Threats to Identification

1.4.2.1 Control Group

The coefficient of interest β in equation (1) can be written as:

$$\beta = E[Y_{i1} - Y_{i0} | T_v = 1] - E[Y_{i1} - Y_{i0} | T_v = 0] \quad (1.3)$$

where $E[Y_{i1} - Y_{i0} | T_v = 1]$ is the difference between women aged 3-9 and 13-18 in 1998 in treated villages and $E[Y_{i1} - Y_{i0} | T_v = 0]$ is the difference between the same cohorts in control villages. We expect the first difference to be positive for education outcomes because the 9 treated villages ($T_v = 1$) got schools after 1998 which benefited girls of the young cohort. We also consider

separately, as a different treatment group, villages that had a nearby school less before 1998 and got a new one after (5 villages). The control group ($T_v = 0$) includes 2 types of villages : (i) those that did not have a school before 1998 and did not benefit from one in 1998-2005 (3 villages) as well as (ii) those that had a nearby school before 1998 and that did not receive a new school after 1998 (13 villages). Among the second type of control villages, 6 were exposed to a nearby school construction between 1988 and 1997. This could lead to an under-estimation bias of β since younger cohorts in these villages were exposed to more schools than older ones thus increasing the value of the second difference $E[Y_{i1} - Y_{i0} | T_v = 0]$. To reduce this bias, we also estimate equation excluding these 6 villages. In these regressions, the control villages ($T_v = 0$) contain villages that are not exposed to nearby school constructions over the period (3 villages) as well as villages that had a nearby school before 1988 and did not get a new school after (7 villages). Keeping seven villages exposed to a nearby school before 1988 in the control group does not seem to be a concern to the identification given that both cohorts aged 3-9 and 13-18 years old in 1998 were equally exposed to these schools. The results presented in Section 5 show that the magnitude of β becomes larger and more statistically significant when excluding the six villages which is consistent with the presence of an under-estimation bias.

1.4.2.2 Sample Selection

As shown in Table 2.1, out-migration is very common in the Niakhar area, 54% of women born in the area move-out. Once individuals leave the area for more than 6 months, they are not followed anymore by the Niakhar HDSS (Lalou and Delaunay 2015) explaining the fact that fertility, marital, and reproductive health behavior data for non-migrant women are used in this paper.

Although the out-migration rate is high in the Niakhar area, it could be really a concern for the identification if out-migration is selective. For instance, if migrants are systematically more likely to be exposed to the sharp increase in school construction during the period 1998-2005 in the Niakhar area and more educated compared to non migrant, this could bias the point estimates for early fertility, child marriage, and health behavior outcomes. For instance, women exposed to the program are more educated and have a willingness to out-migrate from their rural area to urban area looking for job opportunities. To examine whether out-migration is selective, we assess the impact of the treatment on education outcome for both non migrant and migrant women and assess to what extent exposure to the school construction program is correlated with the probability of out-migration. Panel A of Table 1.3 shows that exposure to the school construction program during the period 1998-2005 increases the probability to complete primary

school by 0.16 for non migrant women aged 3-9 in 1998 but has no effect on migrant women aged 3-9 in 1998. When we exclude the six villages exposed to a nearby between 1988 and 1997 in the control group in Panel B of Table 1.3, the magnitude of coefficients becomes larger confirming that non migrant women seem to be more educated than migrant women. Furthermore, Table 1.4 presents the effect of the treatment on the probability of out-migration. Exposure to the school construction program during the period 1998-2005 in the Niakhar area is not correlated with the probability of out-migration. The point estimates is small and statistically insignificant. Tables 1.3 and 1.4 provide evidence suggesting that non migrant women are more educated than migrant women, and exposure to the sharp increase in school construction during the period 1998-2005 is orthogonal to the probability of out-migration. Therefore, we can argue that out-migration does not seem to be a concern for the identification strategy. Given that out-migration is orthogonal to the school construction program during the period 1998-2005 in the Niakhar area, it will simply attenuate the point estimates.

Another potential issue for the identification strategy is internal mobility within the Niakhar area. For instance, the parent's perception on the importance of children's schooling could lead households with a strong preference for education to move closer to primary schools. This might induce a selection bias since children of "movers" would be different compared to non-moving children due to specific unobservable characteristics. We show in Section 6 that the internal mobility rate of children in primary schooling age within the study area between 1998 and 2005 is very low, about 2%, and the magnitude of point estimates remain relatively stable when we exclude internal movers.

1.5 Results

1.5.1 Impact on schooling

1.5.1.1 Main results

Before assessing the impact of the increase in school construction between 1998 and 2005 in the Niakhar area (henceforth called "the program") on early fertility, child marriage, and health behavior outcomes, we focus on education as it is likely to be one of the main pathways which could be at work. As stated above, we consider two treatments separately: First if the village of birth benefited from a school construction after 1998 (nine villages), and second if the village of birth benefited from a new school construction after 1998 (five villages). Results tables are also split between two panels: Panel A presents the point estimates when the sample includes all

thirty villages of the study area, and Panel B displays the results when we exclude six villages that got a school between 1988 and 1997 from the control group (see discussion in the previous section). We use a dummy variable indicating whether women complete primary school as an educational outcome rather than the number of years of schooling because the survey collects the educational level attained by individuals, not the number of years they spent in school.

Table 1.3 presents estimates of the impact of the program on the probability of completing primary school. The first two columns show estimates for non-migrant women, the third and fourth columns refer to estimates for migrant women, and the fifth and sixth columns show estimates for the pooled sample. In all regressions, we control for the village of birth and year of birth through fixed effects, individual characteristics such as religion and caste, and the interaction between the year of birth and 3 village-level variables measured in 1997: the population of schooling age children, their schooling rate, and the population density. Column (1) in panel A of Table 1.3 indicates that the program raises education attainment for non migrant women in villages exposed to a nearby school after 1998. More precisely, exposure to the program increases the probability to complete primary school by 15.7 percentage points for non migrant women exposed to the program. This corresponds to a 36% increase in the probability of completing primary school relative to the mean in the control group. When the six villages are excluded from the control group, the magnitude of the coefficient becomes larger in magnitude and more statistically significant as shown in Panel B of Table 1.3 : the point estimate increases to 23.8 percentage points. Pooling non migrant and migrant women together (column (5) of Panel B), we find that exposure to the program increases the probability to complete primary school by 10.4 percentage points for all treated women compared to women not exposed to the program. Given that migrant women exposed to the program did not experience an increase in educational attainment (column (3)), the point estimate for all women in column (5) of Panel B is driven by non migrant women. Furthermore, we do not find an effect of the program for the second treatment (columns (2) , (4) and (6)). These results suggest that more girls switch to schooling in villages not exposed to a nearby school before 1998 that received a school between 1998 and 2005 in contrast to villages already exposed to a nearby school before 1998 and that benefited a new school construction during the period 1998-2005. Why only non migrant women are affected by the program and not migrant women while both were exposed to the program is an interesting question. The average age of out-migration is 15 years old indicating that migrant women are in the Niakhar area during their primary schooling age before they leave their rural homeland. One likely explanation is the parental behavior of migrant women on the perception of schooling. For instance, parents of migrant women might decide to not enroll their daughter in primary school

due to traditional norms or many other reasons. This question is beyond the scope of the paper.

1.5.1.2 Heterogeneity by cohort of birth

To shed light on the heterogeneity of the effect of the program across cohorts of births, we estimate its effect on each cohort aged from 3 to 12 years old in 1998. Figure A1 presents the estimates from equation (2) for all women (both non migrant and migrant women) with the exclusion of six villages located less than 2 km between 1988 and 1997 in the control group to limit the underestimation bias. The omitted group is women aged between 13 and 22 in 1998. We control for the village of birth and year of birth fixed effect, religion, caste, and three interaction variables as in equation (1). Each dot represents the coefficient of the interaction β_k between a dummy indicating the age in 1998 and a dummy for being born in a treated village.

Figure A1 shows that the youngest cohort in 1998 benefited more from the program relative to the older cohort. The probability to complete primary school is positive for women aged 3-10 in 1998, and statistically significant for women aged 3, 5, and 7 years old in 1998. The magnitude of the effect is higher for women aged 3 in 1998 who experienced 25.7 percentage points increase in the probability to complete primary school compared to women in the control group. The point estimates presented in Figure A1 reinforce our choice of the cohort exposed to the program which is women aged 3-9 in 1998.

1.5.1.3 Testing for the parallel trend assumption

To test the validity of the identification assumption of the difference-in-differences method stressed above, we estimate a placebo test by comparing women born between 1979 and 1984 (aged 14-19 in 1998) with women born between 1973 and 1978 (aged 20-25 in 1998). Both cohorts were no longer of primary school age in 1998 and should not be exposed to the program. Furthermore, the placebo test allows us to make sure that there were no pre-existing differences in educational attainment between treated and control villages prior to the program. For instance, if the educational attainment of women raised faster in treated villages before 1998 compared to control villages, the point estimates in Table 1.3 simply reflect pre-existence trends and capture a spurious effect of the program.

Table A1 in the appendix presents the estimates of the placebo test. In Panel A with all thirty villages of the study area, we find no evidence of pre-existence differences in educational attainment between treated and control villages when we compare two cohorts not exposed to the program. However, in Panel B, when we exclude six villages located less than 2 km to school

during the period 1988-1997 in the control group, we find a negative and significant difference in educational attainment for non migrant. This result indicates that non migrant women aged 20-25 in 1998 in control villages were more educated relative to non migrant women aged 14-19 in 1998 in treated villages. Given that control villages contain villages located less than 2 km to school before 1988 and villages never located less than 2 km to school, the older cohort aged 20-25 in 1998 had access to primary school, in contrast to the younger cohort aged 14-19 in 1998 which could explain the negative and significant difference in the probability to complete primary school in column (1) of Panel B of Table A1. Evidence on the existence of pre-existence trends in educational attainment in the placebo test leads us to interpret the point estimate for non migrant women in column (1) of Panel B of Table 1.3 as a catch-up in educational attainment. In fact, before the program, the non migrant in control villages were on average more educated compared to non migrant in treated villages. After the implementation of the program, the latter group becomes more educated relative to the former group suggesting a catch-up in educational attainment.

1.5.2 Impact on early fertility, child marriage, and health behavior

1.5.2.1 Main results

Table 1.5 shows the estimates of the impact of the program on education and demographic outcomes and reproductive health behavior. The outcomes are ordered as follows: (1) number of births experienced before age 25, (2) probability of being married before age 18, (3) probability to attend at least four prenatal visits at the first birth, (4) probability of delivering at home. As stressed above, the Senegalese Ministry of Health recommends women to attend at least four prenatal visits before delivery explaining our choice to use a dummy variable indicating whether a woman attended at least four prenatal visits at the first birth as an outcome of health behavior. In all specifications, we control for the village of birth and year of birth through fixed effects, individual characteristics such as religion and caste, and the interaction between the year of birth and 3 village-level variables measured in 1997: the population of schooling age children, their schooling rate, and the population density.

In both Panels of Table 1.5, we find no effect of the second treatment on outcomes of interest as in Table 1.3 for the schooling outcome. As explained above, more girls switch to schooling in villages not exposed to a nearby school before 1998 in contrast to villages already exposed before 1998. In Panel A of Table 1.5 with all thirty villages of the Niakhar area, results suggest that

exposure to the program seems to reduce early fertility, child marriage, and delivery at home at the first birth although only the effect on child marriage is statistically significant at 10% level. In addition, exposure to the program is positively correlated to the probability of attending at least 4 prenatal visits at the first birth but the magnitude of the coefficient is small and close to zero. As stressed earlier, these results are likely to be underestimated because of the existence of villages that had access to school in the control group. To limit the underestimation bias, we exclude six villages exposed to a school between 1988 and 1997 in the control group in Panel B of Table 1.5. Once we do so, the point estimates become larger in magnitude and more statistically significant in Panel B confirming the existence of an underestimation bias in Panel A. In Panel B, we find that exposure to the program reduces the number of kids before age 25 by 0.52, the probability of being married before age 18 by 16.2 percentage points, and the probability of delivering at home at the first birth by 14 percentage points. Relative to the mean in the control group, these results represent a 39% decline in early fertility, a 73% decrease in child marriage, and a 43% reduction in delivering at home at the first birth. Our results thus suggest that the impact of the program on outcomes of interest is big. Regarding the existing evidence in the economics literature, our results seem to be in line with Osili and Long (2008) who assessed the impact of a massive school construction program during the period 1976-1981 in Nigeria the most populated country in Africa. The authors find that exposure to the program reduces the number of kids before age 25 by 1.09, which corresponds to a 46% decrease in early fertility relative to the mean (2.35 children per woman aged less than 25 years old). Furthermore, the trend of fertility rate during the period 1984-2017 in the Niakhar area for women aged 15-19 and 20-24 presented in Figure A2 in the appendix highlights that early fertility rate across younger women decreased dramatically while fertility rate among women aged more than 30 remains relatively stable over time. In addition, Figure 1.3 shows that women tend to delay their first marriage, and Figure 1.5 displays the decline of delivering at home over the last three decades. The proportion of women who attended at least four prenatal visits before giving their first birth remains low in the Niakhar area, on average 11% explaining the fact that the effect of the program on prenatal visits is close to zero and statistically insignificant.

1.5.2.2 Testing for the parallel trend assumption

Table A2 in the appendix presents the results of the placebo test in which we compare women born during the period 1979-1984 (aged 14-19 in 1998) with women born between 1973 and 1978 (aged 20-25 in 1998). Both cohorts were not exposed to the program and the point estimates should be smaller in magnitude and statistically insignificant. As expected, we find no

evidence of pre-existing differences in early fertility, child marriage, and health behavior outcomes between treated and control villages when we compare two cohorts not exposed to the program. Thus, our results in Table 1.5 do not capture a continuation of pre-existing trends in demographic and health behavior outcomes before 1998.

1.6 Channels and Discussion

Results presented in the previous section point out two main pathways through which schooling may reduce early fertility in the Niakhar area. First, the decline in early childbearing is driven by the postponement of age at marriage. Exposure to the sharp increase in school construction during the period 1998-2005 decreases the probability of being married before age 18 by 16.2 percentage points. This result is in line with (Keats 2018; Adu Boahen and Yamauchi 2018; Zenebe Gebre 2020). Increased school infrastructures since 1998 in the study area may also change traditional norms on child marriage which could lead parents to delay the marriage of their daughter. Second, change in childbirth location seems to play a key role in the reduction of early fertility. Women exposed to the school construction between 1998 and 2005 are less likely to deliver at home compared to women not exposed. Our interpretation is that schooling allows women to be aware and better understand the risk of maternal and infant mortality of giving birth at home leading to delivery in health facilities rather than home. Using demographic data of the Niakhar area, Bousmah (2017) has shown that child mortality is positively correlated to fertility with a strong effect among women who start childbearing at age 15. Delivering less at home may reduce infant mortality which could prevent women to engage in a catch-up fertility strategy and lowers early fertility. A catch-up fertility strategy also called a replacement behavior means that women tend to replace kids who died by trying to get additional kids quickly. Furthermore, the internal mobility flows within the Niakhar could be a concern for the identification strategy due to the likely selection bias. Figure A6 in the appendix presents the internal mobility rate of children aged between 6 and 11 years old in the Niakhar during the period 1998-2005. Results show that the mobility rate is very low in the study area, about 2% during that period. Table 6 reports the benchmark results and the point estimates when we drop out and indicates that our main findings are relatively stable once we exclude internal movers.

One may argue that husband's education could matter in the fertility behavior of the household. For instance, women exposed to the program have high educational attainment and may choose a partner who is as well educated suggesting that the identification strategy may reflect an overestimation bias of the impact of the program. As shown in Figure 1.3, the age differ-

ence between husbands and spouses is eight years on average in the Niakhar area indicating that husbands were no longer in primary school age in 1998. This large age difference between partners indicates that there is very unlikely that husbands benefited from school construction during the period 1998-2005. Indeed, 88% of husbands were aged more than 12 years old in 1998 suggesting that they were too older to attend primary school, and only 4.5% were aged between 3 and 9 years old in 1998. Husband's schooling does not appear to be a channel at work in the area. Furthermore, contraceptive use and labor market participation could be pathways through which schooling may reduce early fertility, but we do not have data allowing us to test empirically these channels. According to the 2017 DHS data, 27% of women aged between 15 and 49 years old in the Fatick district use contraceptives (ANSD 2018) indicating that the prevalence of contraceptive methods still relatively low compared to urban areas where the percentage is 37%. When we account for the health infrastructures by interacting the treatment variable with the distance to these facilities, we find that the coefficients on the interaction term are positive but not statistically significant, as shown in Table A3 in the appendix. The positive sign of the interaction term suggests that the decline in early fertility, child marriage, and health behavior outcomes induced by the exposure to the school construction is lower as the distance between villages and health facilities increases.

The Niakhar area experienced a meningitis outbreak in 1998-1999 which increased dramatically the mortality rate of children under five years old. As shown in Figure A5 in the appendix, the child mortality rate reached 169‰ and 127‰ in 1998 and 1999, respectively. Given that our definition of the treated cohort is women aged between 3 and 9 years old in 1998, one may worry about a likely composition effect of the treated group. The treated cohort was exposed to the epidemic outbreak while the control cohort aged 13-18 years old in 1998 were not exposed to the shock. The composition effect means that treated women who survived the meningitis outbreak might have specific characteristics observable and unobservable explaining their survival to the shock and those women are very different from the control cohort. To examine whether women exposed to the school construction between 1998 and 2005 are more likely to die of meningitis compared to women not exposed to school construction, we estimate the effect of the treatment on the probability to die due to the meningitis outbreak in 1998-1999. Table 1.7 shows that exposure to the program does not affect the probability to die of meningitis suggesting that a composition effect does not appear to operate. Last, to examine to what extent the secondary schools influence our results, we interact the treatment variable with the distance to secondary schools. There are four high schools in the Niakhar area constructed in 2001, 2005, 2012, 2015 and located outside the nine treated villages. Table A4 in the appendix shows the estimates

when we account for the distance to high schools. The coefficients on interaction terms are statistically insignificant in all specifications suggesting that our results do not seem to be driven by the proximity to secondary schools.

1.7 Conclusion

In this paper, we assess the impact of school construction between 1998 and 2005 in Niakhar, rural Senegal, on educational, demographic and reproductive health behavior outcomes. Exploiting both geographical and temporal variations in school construction, we use a difference-in-differences method. Our results suggest that school infrastructures raise the educational attainment of rural women. We also find that women exposed to school construction during that period have fewer kids compared to women not exposed. The decline in early fertility is driven by the postponement of age at marriage and reproductive health behavioral change related to childbirth location. Women exposed to the program are less likely to deliver at home suggesting that they are aware of maternal and infant mortality risks of giving birth at home. One challenge for the identification strategy is related to out-migration and internal mobility in the study area. First, we provide evidence suggesting that out-migration is orthogonal to the sharp increase in school construction during the period 1998-2005 and does not bias the identification even if the out-migration is high. Second, the internal mobility rate within the study area is low, and the magnitude of the point estimates remains relatively stable when we exclude internal movers. In addition, health facilities do not seem to drive our results when we account for the distance to health infrastructures. However, one limitation of the paper is that we do not have data on contraceptive use and labor market participation to test empirically whether these mechanisms could be at work.

Given that a number of children do not have yet access to primary schools closer to their village in some remote areas in rural Senegal, development policymakers should increase school supply in these areas to reach the Sustainable Development Goal 4 which aims to “ensure inclusive and equitable quality education for all” by 2030. Universal access to school for all, especially in rural areas could lead to accelerating the demographic transition. An interesting follow-up of this study could be to investigate the impact of the sharp increase in school construction during the period 1998-2005 on complete fertility once all exposed women will achieve their overall fertility. In addition, it will be interesting to look at the intergenerational effect of the exposition by examining whether the children of exposed women have better schooling outcomes compared to children of non-exposed women.

REFERENCES

- Adjamagbo, Agnès and Valérie Delaunay (2018). “La reproduction des familles en contexte de changement socioéconomique et culturel In: *Niakhar, mémoires et perspectives: Recherches pluridisciplinaires sur le changement en Afrique* Marseille et Dakar, Éditions IRD et Harmattan Sénégal, 535p”.
- Adu Boahen, Emmanuel and Chikako Yamauchi (2018). “The effect of female education on adolescent fertility and early marriage: evidence from free compulsory universal basic education in Ghana”. *Journal of African Economies* 27.2, pp. 227–248.
- ANSD (2012). *Enquête Démographique et de Santé à Indicateurs Multiples au Sénégal (EDS-MICS) 2010-2011*.
- (2016). *Situation Economique et Sociale de la Région de Fatick*.
- (2018). *Enquête Démographique et de Santé Continue au Sénégal (EDS-Continue 2017)*.
- Bousmah, Marwân-al-Qays (2017). “The effect of child mortality on fertility behaviors is non-linear: new evidence from Senegal”. *Review of Economics of the Household* 15, pp. 93–113.
- Breierova, Lucia and Esther Duflo (2004). *The impact of education on fertility and child mortality: Do fathers really matter less than mothers?* Tech. rep. National bureau of economic research.
- Chari, AV et al. (2017). “The causal effect of maternal age at marriage on child wellbeing: Evidence from India”. *Journal of Development Economics* 127, pp. 42–55.
- Chicoine, Luke (2020). “Free primary education, fertility, and women’s access to the labor market: Evidence from ethiopia”. *The World Bank Economic Review*.
- Delaunay, Valerie et al. (2013). “Profile: the Niakhar health and demographic surveillance system”. *International Journal of Epidemiology* 42.4, pp. 1002–1011.

- Diagne, Abdoulaye (2012). *Evaluation du Programme Décennal de l'Éducation et de la Formation (PDEF) 2000-2011*. CRES.
- Duflo, Esther (2001). "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment". *American Economic Review* 91.4, pp. 795–813.
- Duflo, Esther, Pascaline Dupas, and Michael Kremer (2015). "Education, HIV, and early fertility: Experimental evidence from Kenya". *American Economic Review* 105.9, pp. 2757–97.
- Evans, David K and Amina Mendez Acosta (2021). "Education in Africa: What Are We Learning?" *Journal of African Economies* 30.1, pp. 13–54.
- Garenne, Michel et al. (2018). "Cinquante ans d'évolution de la fécondité à Niakhar (1963-2012) In: *Niakhar, mémoires et perspectives: Recherches pluridisciplinaires sur le changement en Afrique* Marseille et Dakar, Éditions IRD et Harmattan Sénégal, 535p".
- Keats, Anthony (2018). "Women's schooling, fertility, and child health outcomes: Evidence from Uganda's free primary education program". *Journal of Development Economics* 135, pp. 142–159.
- Lalou, Richard and Valérie Delaunay (2015). "Migrations saisonnières et changement climatique en milieu rural sénégalais: forme ou échec de l'adaptation? In: *Les sociétés rurales face aux changements climatiques et environnementaux en Afrique de l'Ouest*. Marseille : IRD, 287-313. (Synthèses). ISBN 978-2-7099-2146-6".
- Lucas, Adrienne M and Isaac M Mbiti (2012). "Access, sorting, and achievement: The short-run effects of free primary education in Kenya". *American Economic Journal: Applied Economics* 4.4, pp. 226–53.
- McGavock, Tamara (2021). "Here waits the bride? The effect of Ethiopia's child marriage law". *Journal of Development Economics* 149, p. 102580.
- Osili, Una Okonkwo and Bridget Terry Long (2008). "Does female schooling reduce fertility? Evidence from Nigeria". *Journal of Development Economics* 87.1, pp. 57–75.
- Ozier, Owen (2018). "The Impact of Secondary Schooling in Kenya: A Regression Discontinuity Analysis". *Journal of Human Resources* 53.1, pp. 157–188.
- PASEC (2004). *Le redoublement: pratiques et conséquences dans l'enseignement primaire au Sénégal 1995-2000. Rapport d'évaluation*.
- Theunynck, Serge (2009). *School construction strategies for universal primary education in Africa: should communities be empowered to build their schools?* World Bank Publications.
- UNESCO (2000). *Dakar Framework for Action: The World Education Forum*.
- UNICEF (2018). *Child Marriage. Latest Trends and Future Prospects*.

WorldBank (2006). *Implementation completion report: quality education for all program (qefa) in support of the first phase of the ten-year education and training program (PDEF)*. World Bank Publications.

Zenebe Gebre, Tihtina (2020). “Free Primary Education, Timing of Fertility, and Total Fertility”. *The World Bank Economic Review* 34.3, pp. 730–748.

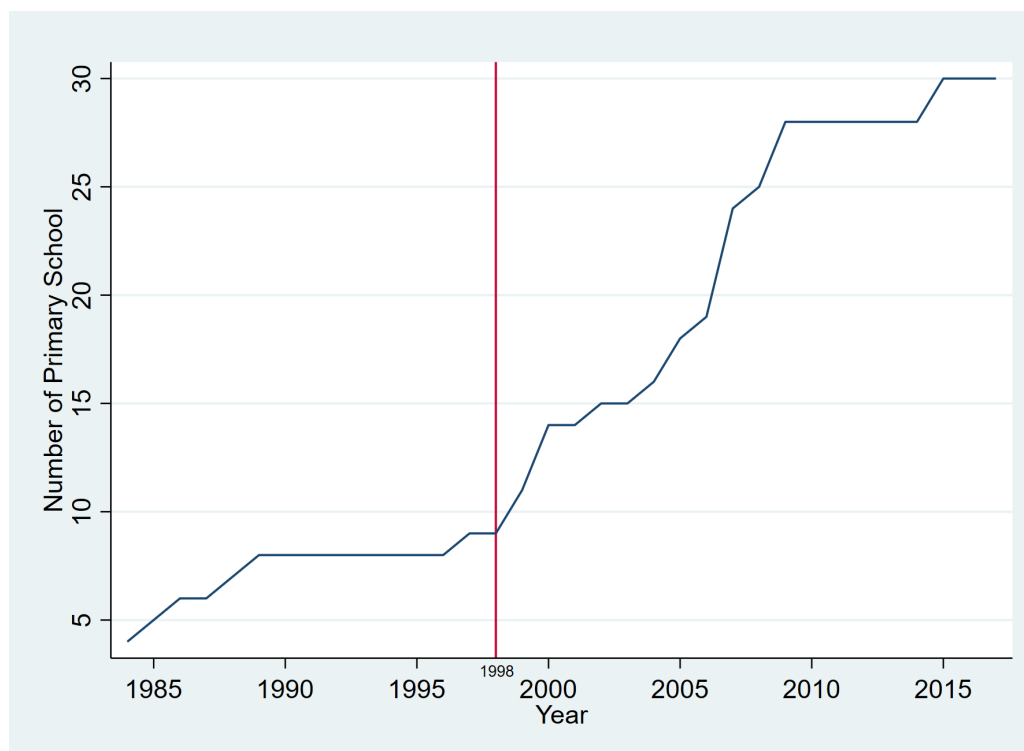


Figure 1.1: Number of Primary school in Niakhar HDSS 1984-2017

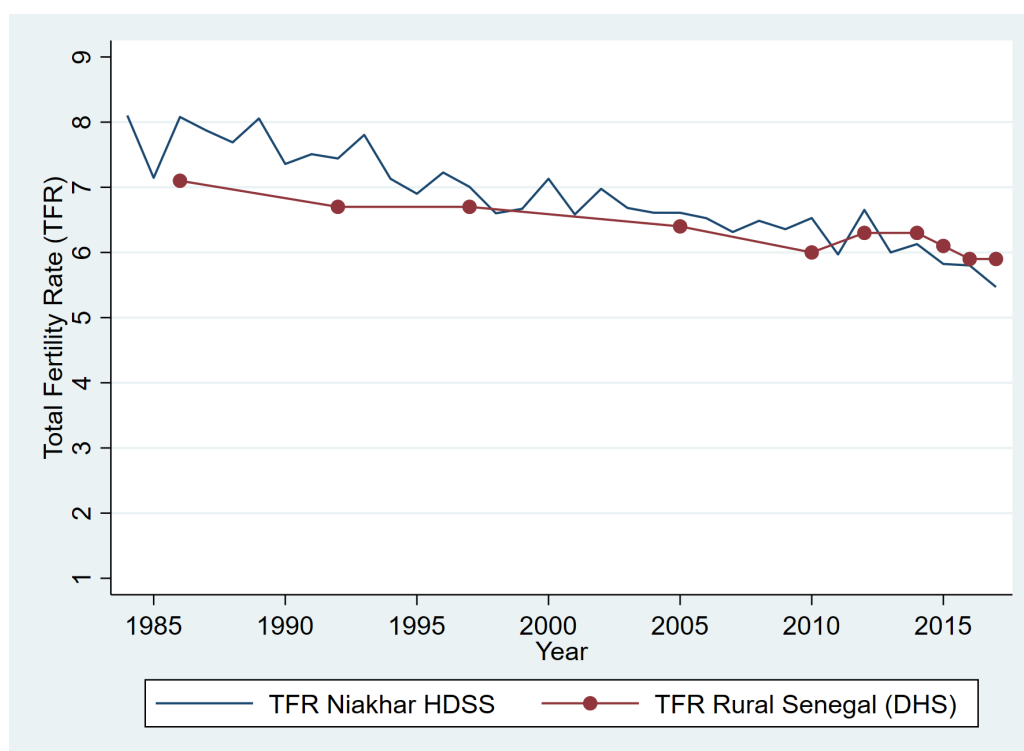


Figure 1.2: Evolution of the Total Fertility Rate (TFR) in Niakhar HDSS and Rural Senegal 1984-2017

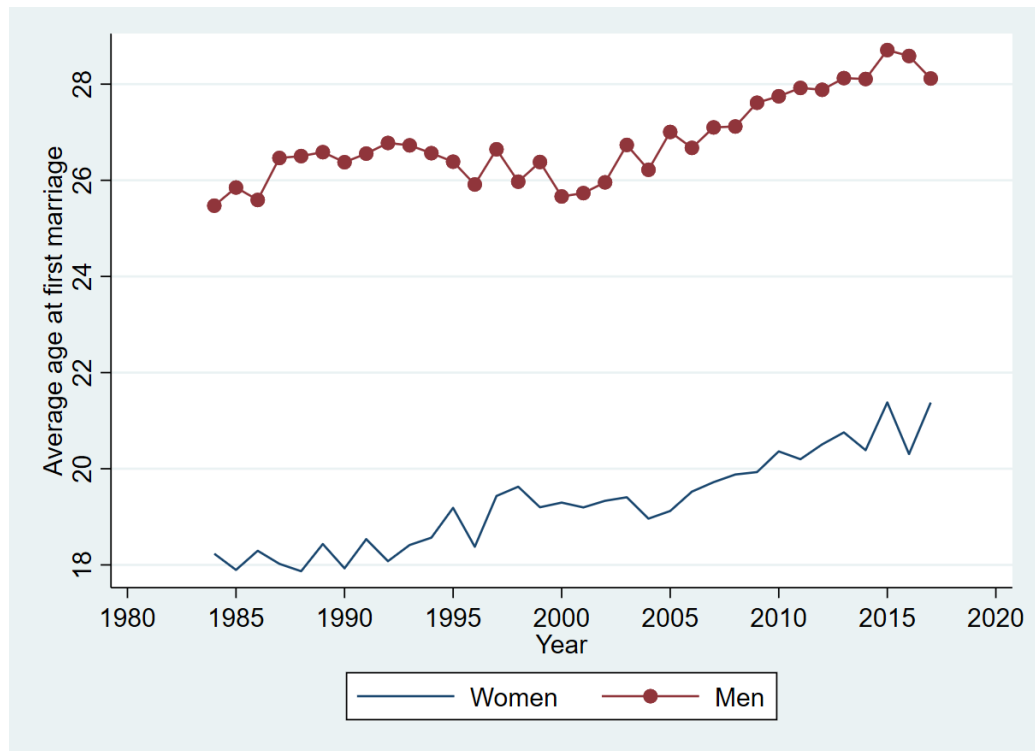


Figure 1.3: Evolution of the age at first marriage in Niakhar HDSS 1984-2017

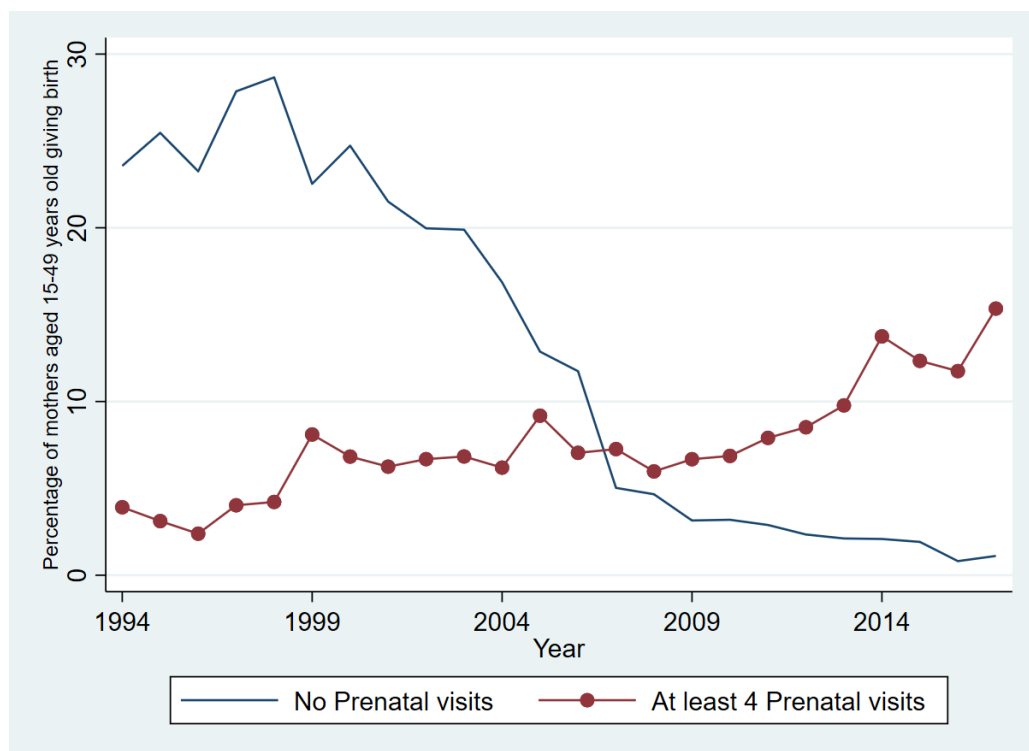


Figure 1.4: Percentage of mothers aged 15-49 years old attended Prenatal visits in Niakhar HDSS 1994-2017

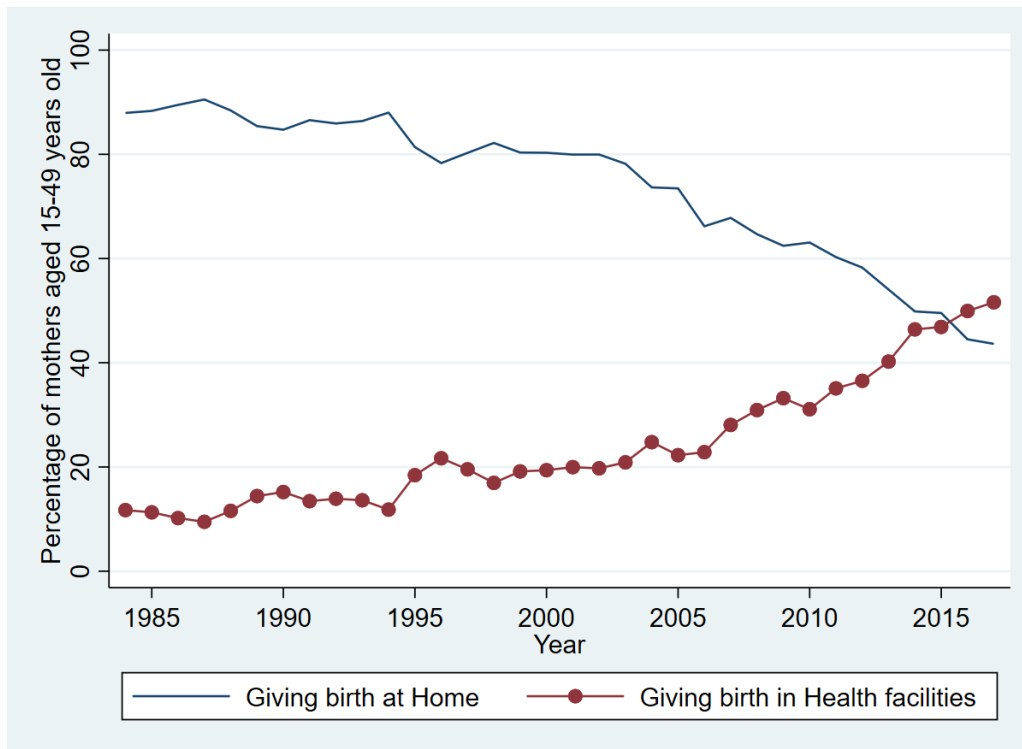


Figure 1.5: Childbirth location of mothers aged 15-49 years old in Niakhar HDSS 1984-2017

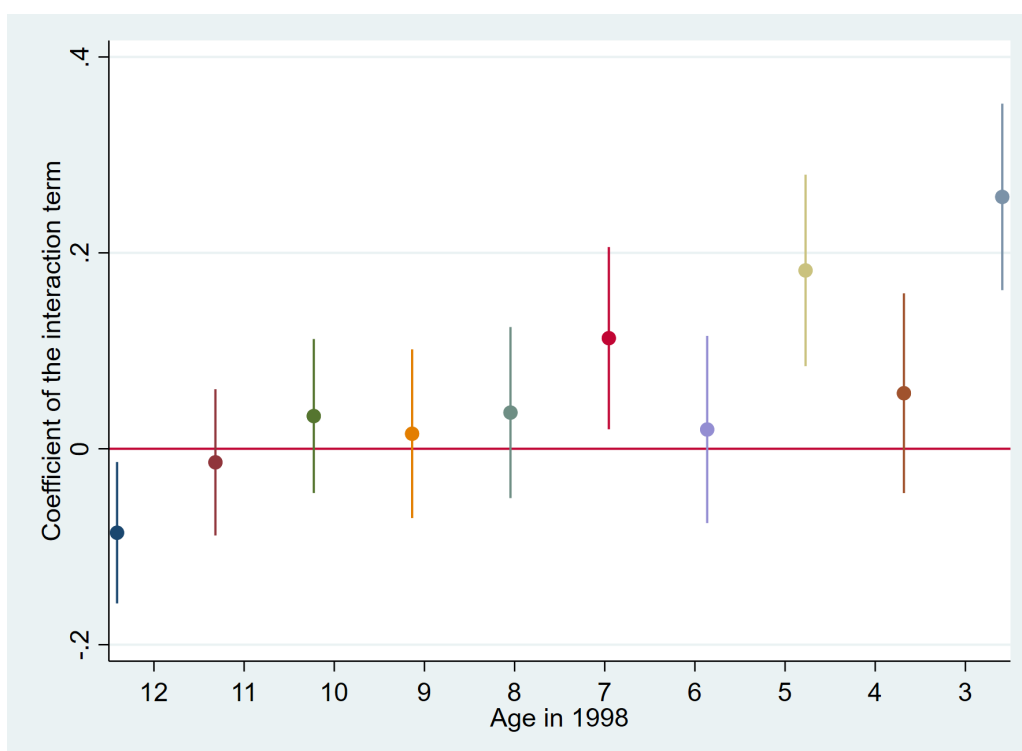


Figure 1.6: Effect of the school construction program on the probability to complete primary school by cohort of birth

Table 1.1: Descriptive Statistics

	Mean	Std. Deviation	Min	Max	N
Age	26.02	(7.99)	15	49	10604
Education					
Attended primary school	0.43	(0.49)	0	1	10604
Completed primary school	0.30	(0.46)	0	1	10604
Fertility-Marriage					
Number of kids per women	1.66	(2.72)	0	15	4919
Number of kids before age 25	0.80	(1.22)	0	5	4919
Married before age 18	0.15	(0.35)	0	1	4919
Age at first birth	20.29	(2.96)	15	36	1923
Age at first Marriage	19.13	(3.88)	12	46	1918
Prenatal care-Childbirth location					
Number of Prenatal visits at first birth	2.47	(1.11)	0	6	1361
At least 4 Prenatal visits at first birth	0.11	(0.31)	0	1	1923
Delivering the first birth at Home	0.51	(0.50)	0	1	1923
Delivering the first birth in Health facilities	0.46	(0.50)	0	1	1923
Migration					
Out-Migration	0.54	(0.50)	0	1	10604
Age at Out-Migration	15.50	(8.92)	0	45	5657
Causes of Out-Migration					
Family reasons	0.37	(0.48)	0	1	5646
Work	0.32	(0.47)	0	1	5646
Marriage	0.18	(0.39)	0	1	5646
Child Fostering for other reasons	0.08	(0.27)	0	1	5646
Child Fostering for schooling	0.02	(0.14)	0	1	5646
Not reported	0.03	(0.17)	0	1	5646
Religion					
Muslim	0.74	(0.44)	0	1	10604
Catholic	0.18	(0.38)	0	1	10604
Other	0.03	(0.16)	0	1	10604
Not reported	0.05	(0.22)	0	1	10604

Table 1.2: Probability to be located less than 2km from the school

	(1)	(2)	(3)	(4)	(5)	(6)
Population of children aged 6-11	0.001*** (0.000)	0.004*** (0.001)	-0.002*** (0.001)			
Population density				0.003*** (0.000)	0.008*** (0.001)	-0.003** (0.001)
Village FE	No	Yes	Yes	No	Yes	Yes
Year FE	No	No	Yes	No	No	Yes
N	660	660	660	660	660	660
R ²	0.154	0.585	0.718	0.074	0.638	0.711

Notes: Village panel (660 observations): 30 villages x 22 years (1984-2005). The dependent variable is a dummy variable equal to 1 if the village of birth was located less than 2 km from the school, and 0 otherwise. Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 1.3: Effect of the school construction program on the probability to complete primary school

	Always Rural		Migrants		All	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All 30 villages						
Less than 2km to school between 1998-2005 * Treat	0.157**		-0.028		0.044	
	(0.066)		(0.038)		(0.036)	
Less than 2km to school before 1998 and between 1998-2005 * Treat		-0.048		-0.039		-0.049
		(0.076)		(0.041)		(0.039)
Mean Dep. Var. (control)	0.43	0.45	0.12	0.13	0.23	0.25
N	1336	1232	2303	2091	3639	3323
Panel B: Excluding 6 villages						
Less than 2km to school between 1998-2005 * Treat	0.238***		-0.004		0.104**	
	(0.073)		(0.044)		(0.042)	
Less than 2km to school before 1998 and between 1998-2005 * Treat		-0.012		-0.016		-0.015
		(0.079)		(0.043)		(0.041)
Mean Dep. Var. (control)	0.40	0.43	0.12	0.13	0.22	0.23
N	1156	1052	2027	1815	3183	2867

Notes: The dependent variable is a dummy equal to 1 if women completed primary school, and 0 otherwise. Treat is a dummy variable equal to 1 if women aged 3-9 in 1998, and 0 if women aged 13-18 in 1998. All specifications include both village of birth and year of birth fixed effects, individual characteristics such as religion and caste, and the interaction between the year of birth and 3 village-level variables measured in 1997: the population of schooling age children, their schooling rate, and the population density. Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 1.4: Effect of the school construction program on the probability of Out-Migration

	Out-Migration	
	(1)	(2)
Panel A: All 30 villages		
Less than 2km to school between 1998-2005 * Treat	-0.040	
	(0.041)	
Less than 2km to school before 1998 and between 1998-2005 * Treat		0.070
		(0.044)
Mean Dep. Var. (control)	0.64	0.63
N	3639	3323
Panel B: Excluding 6 villages		
Less than 2km to school between 1998-2005 * Treat	-0.066	
	(0.048)	
Less than 2km to school before 1998 and between 1998-2005 * Treat		0.059
		(0.046)
Mean Dep. Var. (control)	0.65	0.64
N	3183	2867

Notes: The dependent variable is a dummy equal to 1 if women migrate out of the Niakhar area, and 0 otherwise. Treat is a dummy variable equal to 1 if women aged 3-9 in 1998, and 0 if women aged 13-18 in 1998. All specifications include both village of birth and year of birth fixed effects, individual characteristics such as religion and caste, and the interaction between the year of birth and 3 village-level variables measured in 1997: the population of schooling age children, their schooling rate, and the population density. Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 1.5: Effect of the school construction program on early Fertility, child Marriage, and Health behavioral outcomes

	Number of kids before age 25		Married before age 18		Prenatal visits		Delivered at Home	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All 30 villages								
Less than 2km to school between 1998-2005 * Treat	-0.274 (0.191)		-0.119* (0.061)		0.004 (0.044)		-0.097 (0.066)	
Less than 2km to school before 1998 and between 1998-2005 * Treat		-0.026 (0.215)		0.102 (0.067)		0.048 (0.049)		0.034 (0.072)
Mean Dep. Var. (control)	1.27	1.25	0.20	0.19	0.08	0.09	0.30	0.29
N	1336	1232	1336	1232	1336	1232	1336	1232
Panel B: Excluding 6 villages								
Less than 2km to school between 1998-2005 * Treat	-0.522** (0.213)		-0.162** (0.069)		-0.009 (0.049)		-0.140* (0.073)	
Less than 2km to school before 1998 and between 1998-2005 * Treat		-0.096 (0.225)		0.103 (0.071)		0.045 (0.051)		0.042 (0.075)
Mean Dep. Var. (control)	1.32	1.29	0.22	0.20	0.08	0.08	0.32	0.31
N	1156	1052	1156	1052	1156	1052	1156	1052

Notes: Dependent variables: Number of kids before age 25, Dummy indicating women married before age 18, Dummy indicating women attended at least 4 prenatal visits at first birth, Dummy indicating women delivered at home at first birth. Treat is a dummy variable equal to 1 if women aged 3-9 in 1998, and 0 if women aged 13-18 in 1998. All specifications include both village of birth and year of birth fixed effects, individual characteristics such as religion and caste, and the interaction between the year of birth and 3 village-level variables measured in 1997: the population of schooling age children, their schooling rate, and the population density. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 1.6: Robustness check - excluding internal migrants

	Benchmark results					Excluding internal migrants				
	Educ. (1)	Fert. (2)	Marri. (3)	Prenat. (4)	Deliv. (5)	Educ. (6)	Fert. (7)	Marri. (8)	Prenat. (9)	Deliv. (10)
Panel A: All 30 villages										
Less than 2km to school between 1998-2005 * Treat	0.157** (0.066)	-0.274 (0.191)	-0.119* (0.061)	0.004 (0.044)	-0.097 (0.066)	0.159** (0.068)	-0.276 (0.195)	-0.129** (0.061)	-0.005 (0.046)	-0.138** (0.067)
Mean Dep. Var. (control)	0.43	1.27	0.20	0.08	0.30	0.43	1.26	0.19	0.09	0.30
N	1336	1336	1336	1336	1336	1273	1273	1273	1273	1273
Panel B: Excluding 6 villages										
Less than 2km to school between 1998-2005 * Treat	0.238*** (0.073)	-0.522** (0.213)	-0.162** (0.069)	-0.009 (0.049)	-0.140* (0.073)	0.229*** (0.076)	-0.471** (0.218)	-0.161** (0.070)	-0.016 (0.050)	-0.162** (0.075)
Mean Dep. Var. (control)	0.40	1.32	0.22	0.08	0.32	0.40	1.32	0.21	0.09	0.32
N	1156	1156	1156	1156	1156	1101	1101	1101	1101	1101

Notes: Dependent variables: Dummy indicating women completed primary school in columns (1)-(6), Number of kids before age 25 in columns (2)-(7), Dummy indicating women married before age 18 in columns (3)-(8), Dummy indicating women attended at least 4 prenatal visits at first birth in columns (4)-(9), Dummy indicating women delivered at home at first birth in columns (5)-(10). Treat is a dummy variable equal to 1 if women aged 3-9 in 1998, and 0 if women aged 13-18 in 1998. All specifications include both village of birth and year of birth fixed effects, individual characteristics such as religion and caste, and the interaction between the year of birth and 3 village-level variables measured in 1997: the population of schooling age children, their schooling rate, and the population density. Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 1.7: Exposure to school construction program and Meningitis outbreak 1998-1999

	Meningitis outbreak
	(1)
Panel A: All 30 villages	
Less than 2km to school between 1998-2005 * Treat	-0.0028 (0.010)
Mean Dep. Var. (control)	0.015
N	3694
Panel B: Excluding 6 villages	
Less than 2km to school between 1998-2005 * Treat	0.0001 (0.012)
Mean Dep. Var. (control)	0.014
N	3231

Notes: The dependent variable is a dummy indicating whether a woman died due to the meningitis outbreak. Treat is a dummy variable equal to 1 if women aged 3-9 in 1998, and 0 if women aged 13-18 in 1998. All specifications include both village of birth and year of birth fixed effects, individual characteristics such as religion and caste, and the interaction between the year of birth and 3 village-level variables measured in 1997: the population of schooling age children, their schooling rate, and the population density. Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

1.8 Appendix

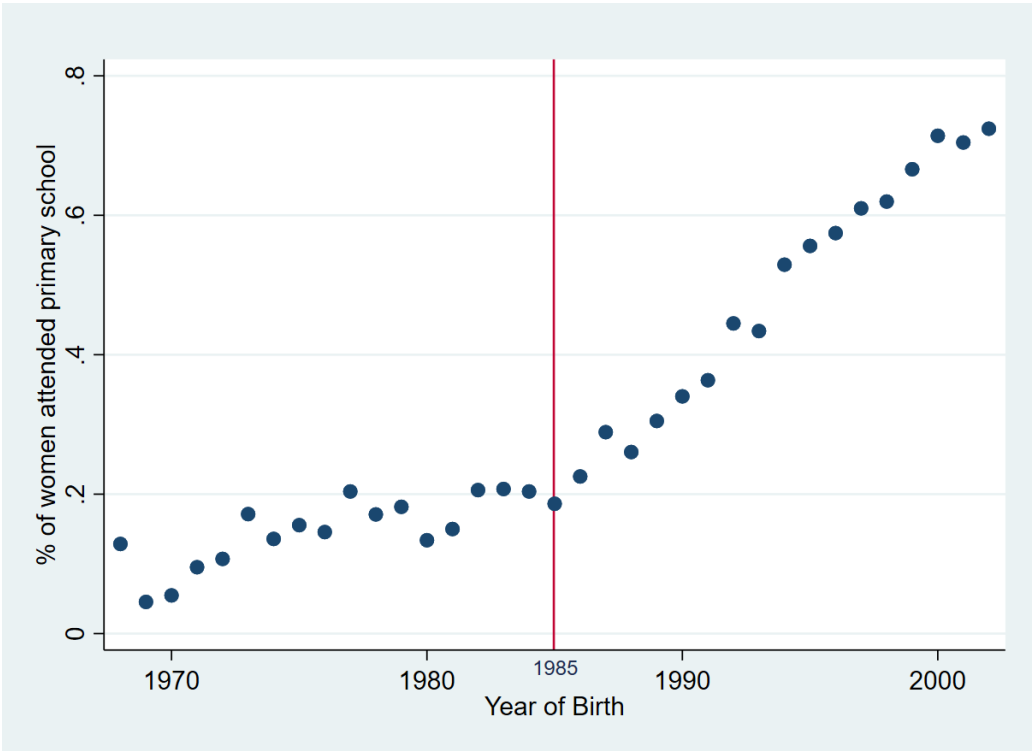


Figure A1: Percentage of women who attended primary school by year of birth in Niakhar HDSS

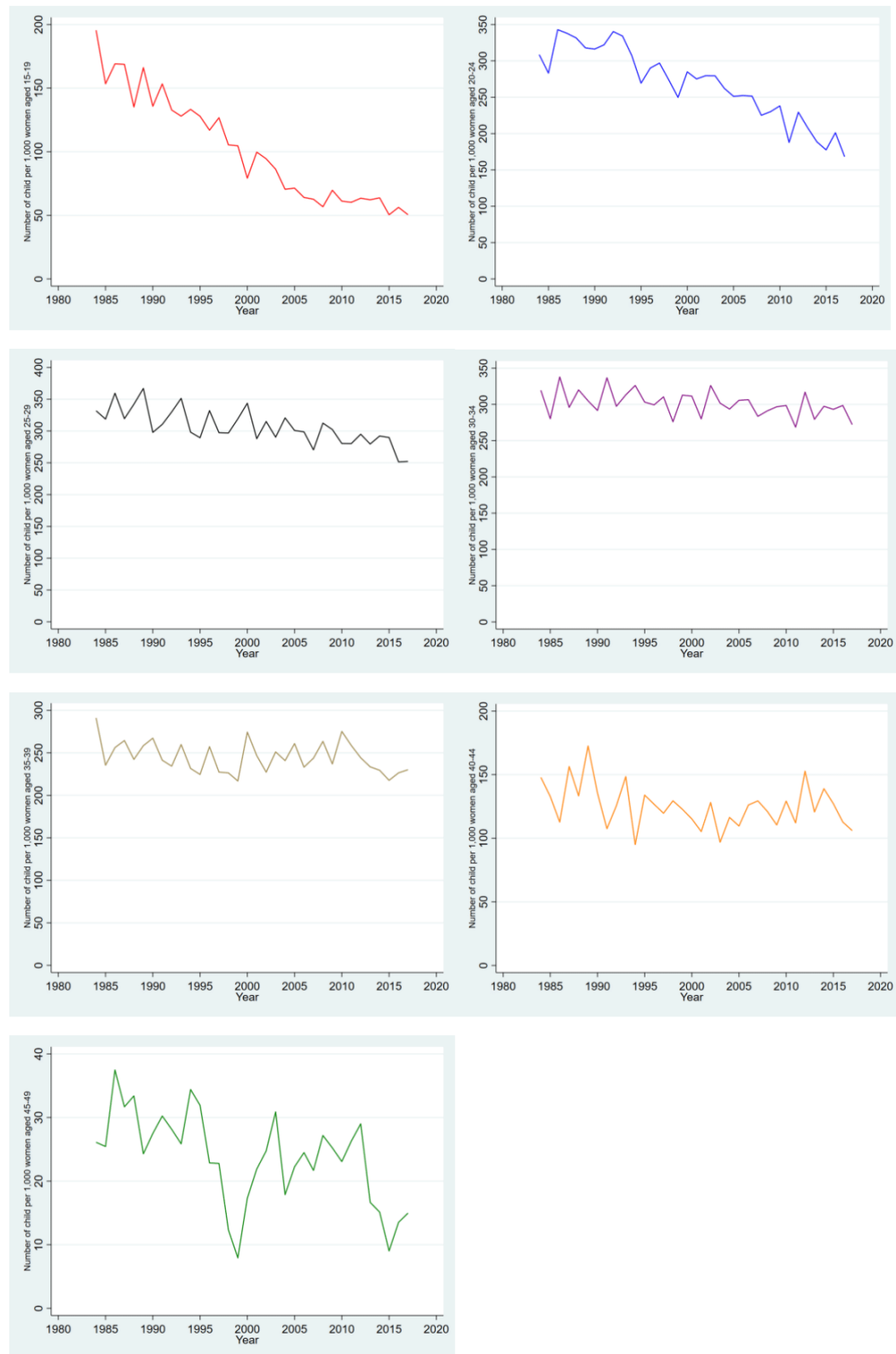


Figure A2: Fertility rate by cohort in Niakhar HDSS 1984-2017

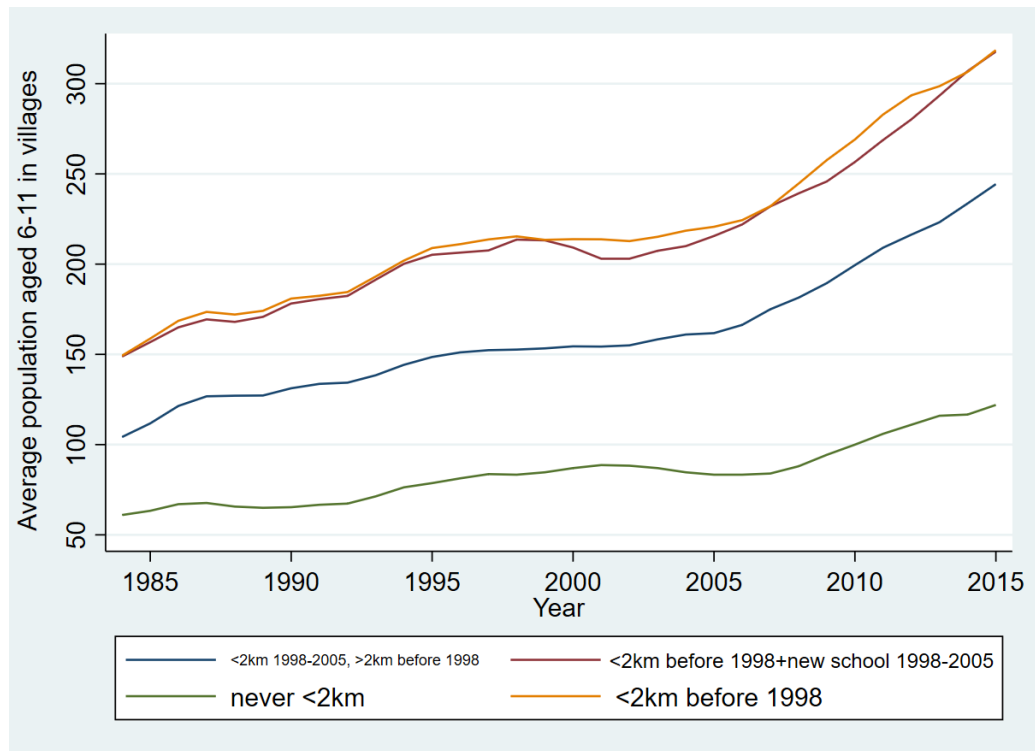


Figure A3: Population aged 6-11 years old in treated and control villages

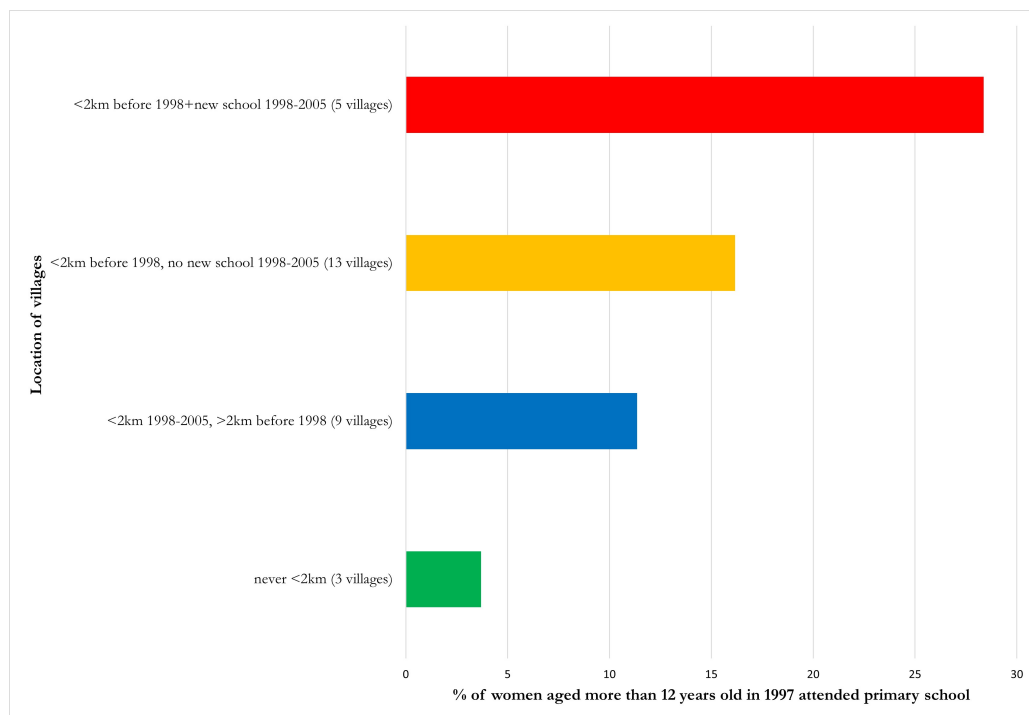


Figure A4: Primary school attendance rate in 1997 (before the sharp increase in school construction in Niakhar HDSS)

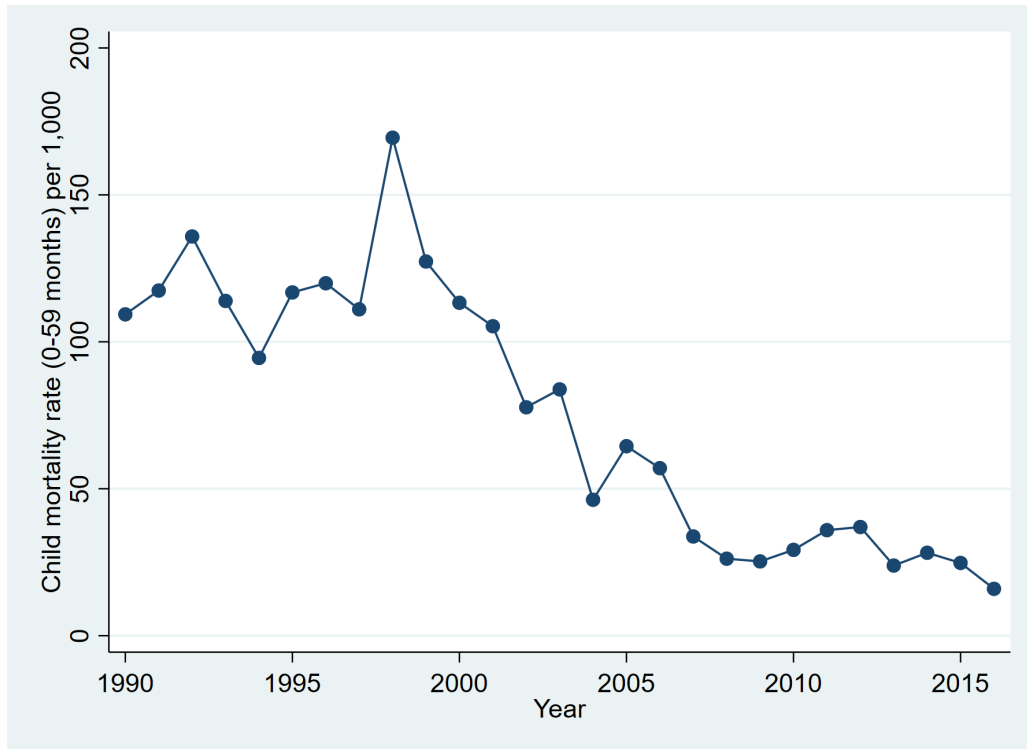


Figure A5: Child mortality rate in the Niakhar area 1990-2016

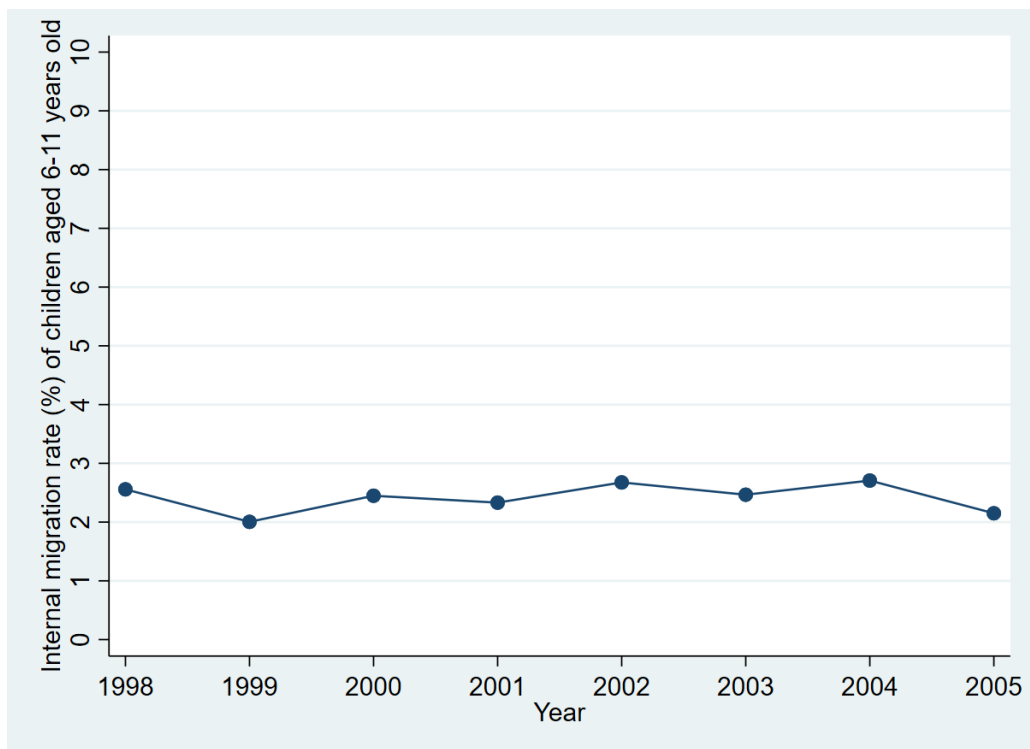


Figure A6: Internal mobility rate within the Niakhar area 1998-2005

Table A1: Effect of the school construction program on the probability to complete primary school - Placebo test

	Always Rural		Migrants		All	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All 30 villages						
Less than 2km to school between 1998-2005 * Placebo	-0.053 (0.051)		0.213 (0.366)		-0.007 (0.038)	
Less than 2km to school before 1998 and between 1998-2005 * Placebo		-0.092 (0.076)		-0.209 (0.255)		-0.036 (0.050)
Mean Dep. Var. (control)	0.06	0.07	0.07	0.08	0.06	0.08
N	516	483	795	622	1311	1105
Panel B: Excluding 6 villages						
Less than 2km to school between 1998-2005 * Placebo	-0.139** (0.060)		0.210 (0.376)		-0.052 (0.046)	
Less than 2km to school before 1998 and between 1998-2005 * Placebo		-0.102 (0.082)		-0.209 (0.265)		-0.031 (0.053)
Mean Dep. Var. (control)	0.06	0.08	0.08	0.09	0.07	0.09
N	459	426	736	563	1195	989

Notes: The dependent variable is a dummy equal to 1 if women completed primary school, and 0 otherwise. Placebo is a dummy variable equal to 1 if women aged 14-19 in 1998, and 0 if women aged 20-25 in 1998. All specifications include both village of birth and year of birth fixed effects, caste, religion, interaction between the year of birth dummies and the population of children aged 6-11 in 1997 in the village of birth, interaction between the year of birth and the share of the population aged 6-11 in 1997 in the village of birth who attended primary school, and interaction between the year of birth and the population density in the village of birth in 1997. Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table A2: Effect of the school construction program on early Fertility, child Marriage, and Behavioral outcomes - Placebo test

	Number of kids before age 25	Married before age 18	Prenatal visits	Delivered at Home
	(1)	(2)	(3)	(4)
Panel A: All 30 villages				
Less than 2km to school between 1998-2005 * Placebo	-0.115 (0.286)	0.103 (0.106)	-0.001 (0.060)	0.019 (0.111)
Mean Dep. Var. (control)	2.15	0.33	0.07	0.65
N	516	516	516	516
Panel B: Excluding 6 villages				
Less than 2km to school between 1998-2005 * Placebo	0.113 (0.328)	0.155 (0.121)	0.004 (0.067)	0.058 (0.130)
Mean Dep. Var. (control)	2.10	0.32	0.06	0.64
N	459	459	459	459

Notes: Dependent variables: Number of kids before age 25, Dummy indicating women married before age 18, Dummy indicating women attended at least 4 prenatal visits at first birth, Dummy indicating women delivered at home at first birth. Placebo is a dummy variable equal to 1 if women aged 14-19 in 1998, and 0 if women aged 20-25 in 1998. All specifications include both village of birth and year of birth fixed effects, caste, religion, interaction between the year of birth dummies and the population of children aged 6-11 in 1997 in the village of birth, interaction between the year of birth and the share of the population aged 6-11 in 1997 in the village of birth who attended primary school, and interaction between the year of birth and the population density in the village of birth in 1997. Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table A3: Heterogeneous effects by distance to Health posts

	Number of kids before age 25	Married before age 18	Prenatal visits	Delivered at Home
	(1)	(2)	(3)	(4)
Panel A: All 30 villages				
Less than 2km to school between 1998-2005 * Treat	-0.558 (0.398)	-0.225* (0.127)	-0.013 (0.092)	-0.233* (0.136)
Less than 2km to school between 1998-2005 * Treat * Distance to Health posts	0.095 (0.117)	0.035 (0.037)	0.006 (0.027)	0.045 (0.040)
Mean Dep. Var. (control)	1.27	0.20	0.08	0.30
N	1336	1336	1336	1336
Panel B: Excluding 6 villages				
Less than 2km to school between 1998-2005 * Treat	-0.642 (0.408)	-0.235* (0.132)	0.001 (0.094)	-0.258* (0.140)
Less than 2km to school between 1998-2005 * Treat * Distance to Health posts	0.041 (0.120)	0.025 (0.039)	-0.003 (0.027)	0.040 (0.041)
Mean Dep. Var. (control)	1.32	0.22	0.08	0.32
N	1156	1156	1156	1156

Notes: Dependent variables: Number of kids before age 25, Dummy indicating women married before age 18, Dummy indicating women attended at least 4 prenatal visits at first birth, Dummy indicating women delivered at home at first birth. Treat is a dummy variable equal to 1 if women aged 3-9 in 1998, and 0 if women aged 13-18 in 1998. All specifications include both village of birth and year of birth fixed effects, caste, religion, interaction between the year of birth dummies and the population of children aged 6-11 in 1997 in the village of birth, interaction between the year of birth and the share of the population aged 6-11 in 1997 in the village of birth who attended primary school, and interaction between the year of birth and the population density in the village of birth in 1997. Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table A4: Heterogeneous effects by distance to Secondary schools

	Number of kids before age 25	Married before age 18	Prenatal visits	Delivered at Home
	(1)	(2)	(3)	(4)
Panel A: All 30 villages				
Less than 2km to school between 1998-2005 * Treat	-0.537 (0.377)	-0.216* (0.120)	0.003 (0.087)	-0.234* (0.129)
Less than 2km to school between 1998-2005 * Treat Distance to Secondary schools	0.093 (0.115)	0.034 (0.037)	0.000 (0.027)	0.049 (0.039)
Mean Dep. Var. (control)	1.27	0.20	0.08	0.30
N	1336	1336	1336	1336
Panel B: Excluding 6 villages				
Less than 2km to school between 1998-2005 * Treat	-0.614 (0.389)	-0.221* (0.126)	0.018 (0.089)	-0.259* (0.133)
Less than 2km to school between 1998-2005 * Treat Distance to Secondary schools	0.034 (0.119)	0.021 (0.038)	-0.010 (0.027)	0.044 (0.041)
Mean Dep. Var. (control)	1.32	0.22	0.08	0.32
N	1156	1156	1156	1156

Notes: Dependent variables: Number of kids before age 25, Dummy indicating women married before age 18, Dummy indicating women attended at least 4 prenatal visits at first birth, Dummy indicating women delivered at home at first birth. Treat is a dummy variable equal to 1 if women aged 3-9 in 1998, and 0 if women aged 13-18 in 1998. All specifications include both village of birth and year of birth fixed effects, caste, religion, interaction between the year of birth dummies and the population of children aged 6-11 in 1997 in the village of birth, interaction between the year of birth and the share of the population aged 6-11 in 1997 in the village of birth who attended primary school, and interaction between the year of birth and the population density in the village of birth in 1997. Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

CHAPTER 2

RAINFALL AND INFANT MORTALITY IN RURAL SENEGAL

Abstract¹

In this paper, we investigate the impact of rainfall shocks on post-neonatal and child mortality in rural Senegal over the period 1985-2016, using a long panel data set and exploiting rainfall variation across communes and over time. Two opposite effects are likely to drive the impact of rainfall on infant mortality: a negative income-nutrition impact and a positive disease-environment impact. Our results show that a positive deviation of rainfall from the historical mean increases the post-neonatal mortality rate. This suggests that the disease-environment effect dominates the income-nutrition effect. Coherently, the data indicates that post-neonatal mortality is driven by malaria and respiratory diseases. Analyzing the heterogeneity of the impact of rainfall across gender, we find significant heterogeneity suggesting that girls but not boys are affected by rainfall shocks. Over the period, the disease-environment effect is concentrated in the first two decades of the observation period during which post-neonatal mortality rate was high. While average rainfall was higher in the last decade, the massive distribution of insecticide treated bed nets allowed keeping malaria mortality at relatively low levels.

JEL Classification: I15, J16, Q54.

Keywords: Rainfall, Infant mortality, Gender, Niakhar HDSS, Rural Senegal.

¹This chapter is co-authored with Anne-Sophie Robilliard. It is under review in the journal *Demography*.

2.1 Introduction

Climate scientists predict that climate change will increase rainfall variability and the occurrence of extreme rainfall events (IPCC 2018). In West Africa, climate model projections indicate that both the total amount of rainfall and the length of dry spells are likely to increase (Sylla et al. 2015; Diedhiou et al. 2018). Although considerable uncertainty remains, this suggests that rainfall events are likely to become more intense. Some studies have shown an increase in precipitations with large spatial and inter-annual variations in Sahel and Senegal over the last two decades (Ali and Lebel 2009; Bodian 2014). These environmental changes and the high frequency of extreme rainfall events impact a wide range of outcomes such as agricultural yields, household incomes and consumption, human capital, health and infant mortality, particularly in the poorest regions of the world (Dell, Jones, and Olken 2014). According to the United Nations Office for Disaster Risk Reduction (UNISDR 2015), low- and middle-income countries have experienced more than 40% of the total number of weather-related disasters that occurred between 1995 and 2015 at a cost of 358,000 lives lost, with about 60% of all deaths recorded during this period.

The detrimental effects of climatic shocks on human health has lead researchers to try and understand how extreme climate events such as droughts, floods, or extreme temperatures affect infant mortality. Evidence on the impact of weather shocks on health outcomes and mortality exists but there is no consensus in the literature about the pathways through which these shocks operate. According to economists, one possibly important channel is the income and nutrition channel (Maccini and Yang 2009; Kim 2010; Burgess et al. 2011; Flatø and Kotsadam 2014; Kumar, Molitor, and Vollmer 2016; Dinkelman 2017; Shah and Steinberg 2017; Hyland and Russ 2019). In developing countries, particularly in rural areas, the majority of the population still derive their livelihoods from farming, which relies to a large extent on rain-fed agriculture. Thus, higher rainfall is likely to increase crop yields, reduce food prices, enhance nutrient intake and consumption and therefore improve health outcomes and lower infant mortality. Furthermore, a number of papers in the economics literature focus on the fetal origins hypothesis which indicates that nutritional deprivation in utero - i.e. the year before birth - may have adverse long-term effects on a wide range of outcomes in adulthood (Almond, Currie, and Duque 2018). Another channel is related to the so-called disease environment. Indeed, a wet year is associated with an intensification of mosquito breeding and is expected to raise the prevalence of malaria, leading to a higher risk of infant mortality (Ndiaye et al. 2001; Kudamatsu, Persson, and Strömberg 2012; WHO 2015). Instead, droughts might reduce water and vector-borne

diseases and lower the risk of death (Henry and Dos Santos 2013; Rabassa, Skoufias, and Jacoby 2014). However, in certain areas, a rainfall shortage can also induce water scarcity and reduce the availability of safe drinkable water thus raising the prevalence of infectious diseases (Bandyopadhyay, Kanji, and Wang 2012; Rocha and Soares 2015).

On the whole, the evidence on the impact of rainfall on infant mortality is mixed and ambiguous and the findings in the literature vary depending on the setting. Furthermore, examining this relationship requires good quality longitudinal data over sufficient long periods of time, and, in developing countries, the analysis is constrained because of the absence of well functioning vital registration systems (Geruso and Spears 2018), especially in rural areas. Given the paucity of administrative data, analysts have turned to using survey data. However, births and deaths are often under reported in poor countries (UNISDR 2015), which may affect the quality of data collected in surveys at the national level and therefore the reliability of research findings.

This paper investigates the impact of rainfall on post-neonatal and child mortality over the period 1985-2016 using data from the Health and Demographic Surveillance System (HDSS) of Niakhar. Located in rural Senegal, the Niakhar is the oldest HDSS in sub-Saharan Africa. Demographic events such as births, deaths, and migration have been recorded at least twice a year since 1962, resulting in a long-term panel and high-quality demographic data which is relatively rare in an African context. In this paper, this data set is combined with observed rainfall data over the period 1985-2016 to measure the impact of rainfall on post-neonatal and child mortality. We also examine the heterogeneous effects of rainfall across gender, season, proximity to health facilities as well as by decades.

This paper contributes in several ways to the growing literature that addresses the causal relationship between climate change and health outcomes. First, the paper uses high-quality longitudinal demographic data collected over a very long period of time. The existing evidence in sub-Saharan Africa relies on Demographic and Health Surveys (DHS), which are subject to recall bias and omission of demographic events, and do not document the children's causes of death. Second, we differentiate the impact between seasons and by gender, as well as by decades. Third, we make use of recorded causes of deaths to support the interpretation of our findings. Causes of death data are not available in most of the studies in the literature and therefore the interpretation of specific channels linking precipitation variability and health outcomes is often tentative.

A number of results stand out. Over the period of study, a 10% increase in the deviation of rainfall from the historical mean during the rainy season raises contemporaneous post-neonatal mortality rate by 1.99 points – i.e. 8.9%. Looking at the causes of death data, we find that

the higher incidence of malaria and respiratory diseases during the rainy season are the main causes of post-neonatal mortality consistent with the disease environment effect. Examining the heterogeneity of the effects, we find a striking gender effect suggesting that girls are affected by rainfall shocks but not boys. Both mosquito nets and infrastructures appear to mitigate the adverse effects of positive rainfall shocks on post-neonatal and child mortality. Analyzing the heterogeneous effects of rainfall by decade, we show that the disease-environment effect is concentrated in the first two decades. During the last decade characterized by increased rainfall levels with important inter-annual variability, the income-nutrition effect dominates the disease-environment effect suggesting that public health programs such as the National Malaria Control Program performed well.

The remainder of the paper is organized as follows. Section 2 reviews the existing literature on the impact of rainfall variability on health outcomes and infant mortality. Section 3 presents the Niakhar HDSS. Section 4 describes the data, the definition of key variables, and presents some summary statistics. Section 5 explains the identification strategy. Section 6 presents the main results, the heterogeneity of results by season, gender, proximity to health facilities, and decade to shed light on the specific channels through which rainfall may affect mortality.

2.2 How do rainfall fluctuations affect health outcomes and infant mortality ?

Investigating the impact of rainfall on child health outcomes is challenging due to the various potential pathways through which this impact is likely to operate. The literature identifies two main channels through which rainfall may affect health outcomes and infant mortality: the income-nutrition effect and the disease environment effect. We present the empirical literature associated with these two channels in turn.

The theoretical rationale for the existence of an income-nutrition effect of rainfall on mortality rests on the observation that rural populations rely extensively on rain-fed agriculture as farmers typically do not have access to irrigation. When rainfall levels drop, crop yields fall, incomes plummet and food prices rise. The production shock leads to a decrease in nutrient intake of children therefore increasing the risk of infant mortality. A number of empirical studies supports the existence of the income-nutrition effect. Using the Indonesian Family Life Survey (2000) linked with district-year of birth rainfall, Maccini and Yang (2009) analyze the long-run effects of rainfall on health outcomes of women and men born between 1953 and 1974 in rural areas. To deal with measurement errors in rainfall data, the authors run an instru-

mental variables estimation in which early-life rainfall is instrumented with four alternative rainfall variables from more distant rain stations. One of the main findings is that a positive rainfall deviation from the historical mean of the district in the year of birth increases the probability of self-reporting a very good health status for adult women. According to the authors, discrimination in the allocation of household resources may explain why rainfall shocks affect the long-run health outcomes of women but not of men. Looking at rural India, both Burgess et al. (2011) and Kumar, Molitor, and Vollmer (2016) show that negative rainfall shocks such as droughts lead to higher mortality rates. Burgess et al. (2011) rely on vital statistics data from 1957 to 2000 combined with temperature and precipitations data at the district level to assess the impact of weather shocks on death. The vital statistics data provide information on births and deaths in India and allow the construction of two outcome variables: the infant mortality rate (under age one) and the overall mortality rate (over age one) at the district level. In order to quantify the impact of rainfall on mortality, the authors construct terciles of total annual precipitations (lower, middle, and upper) at the district level. They find that the occurrence of the lowest tercile – which corresponds to a negative rainfall shock such as drought – increases mortality rates significantly in rural areas. However, the authors point out that mortality rates are likely to be underestimated due to under-reporting of births and deaths: only 55% of births and 46% of deaths were registered in 2000 according to the National Population Commission of India. This highlights the limitations of civil registration systems to accurately measure vital events in developing countries. Kumar, Molitor, and Vollmer (2016) use District Level Household Survey data (DLHS) linked with rainfall data at the district level to investigate the impact of drought experienced in utero (i.e. the year before birth) and in the year of birth on Weight-for-Age z-scores (WAZ) and infant mortality rate in rural India. Drought is defined as a dummy variable that takes the value 1 when total precipitation from June to September is equal to 75 percent of its historical mean. They find that drought exposure in utero or in the year of birth significantly reduces the WAZ by 0.1 standard deviations for children born between June and December. Furthermore, drought exposure in the year of birth raises the infant mortality rate by 10 deaths for 1,000 live births for children born during the second half of the year. Still in India, Shah and Steinberg (2017) analyze the relationship between early-life exposure to rainfall shocks and human capital outcomes in later life. A positive rainfall shock is defined as yearly precipitation above the 80th percentile and a negative shock as precipitations below the 20th percentile within the district. Using the India Human Development Survey (IHDS), the authors find that a positive rainfall shock either during the year of birth or the year before increases the weight-for-age z-scores significantly for children 1 to 5 years old in rural areas due to the higher

nutrient intake and consumption and therefore lead to a greater school investment and higher levels of human capital in later life. In the Sub-Saharan setting, several studies shed light on the income and nutrition effect as a potential channel between rainfall and health outcomes in early life. Using the 1996 South Africa census data combined with rainfall data, Dinkelman (2017) examines the impact of drought exposure from in utero to age four on the probability of having a disability (vision, hearing or speech, mental or physical) in later-life. The author computes the Standardised Precipitation Index (SPI) to determine a dry year. In each district of birth, a year is characterized as dry if the value of the SPI is below -1.5 as defined in the climate literature. Focusing on cohorts born from 1948 to 1986 (aged between 10 to 48 years old in 1996) in rural areas, the author finds that early-life exposure to drought increases the probability of reporting a disability by 3.5%. Negative rainfall shocks affect both men and women significantly but the size of the effect is twice as high for men. According to the author, one way to explain this gender bias is that males are more vulnerable to nutritional and health shocks induced by drought during infancy. Hyland and Russ (2019) rely on Demographic and Health Surveys (DHS) from nineteen Sub-Saharan Africa countries matched with weather data and find that exposure to drought during infancy has a strong negative intergenerational effect on health. Following Dinkelman (2017), extreme drought event is characterized by a Standardised Precipitation Index (SPI) of 2 standard deviations below the historical average. The authors focus on women born between 1953-1998 and define their treatment variable - women's exposure to extreme drought in early-life - as the percentage of years exposed to any extreme drought event between the year of birth to age four. They find that early-life exposure to extreme drought significantly lowers height and wealth for adult women and raises their probability of giving birth to low-weighted children ($< 2.5kg$) by 2.6 percentage points. The pathway through which these effects operate is that extreme droughts induce an income shock and significant nutritional deprivation during infancy which adversely affects physical and cognitive development, lowers years of schooling in later life and therefore leads to worse health and economic conditions in adulthood. Conversely, using DHS data from nine West African countries, Kim (2010) finds that higher rainfall can also adversely affect infant mortality rate. She finds that a positive one standard deviation shock in precipitations raises the probability of death before age one by about 0.6%. Her interpretation is that higher rainfall drives the opportunity cost of labor of mothers up which leads to a reduction in breastfeeding time with negative consequences on infant mortality. Concerning the gender difference of impacts of negative rainfall shocks, Flatø and Kotsadam (2014) examine the relationship between drought and gender bias in infant mortality in sub-Saharan Africa. Using DHS data combined with rainfall data, the authors show that droughts induce the deaths of 12

more girls per 1000 births than boys. According to the authors, this effect operates through an increased allocation of resources and time towards boys during hardship.

Another channel through which rainfall is likely to affect mortality is the so-called disease-environment effect. Indeed, a number of studies stress that this channel plays an important role in developing countries where vector-borne diseases such as malaria are important causes of death. According to the World Health Organization WHO (2015), Africa has the highest prevalence of malaria with about 90% of all malaria deaths worldwide in 2013. Furthermore, 78% of malaria deaths occur in children under 5. Kudamatsu, Persson, and Strömberg (2012) use DHS data from twenty-eight African countries and find that higher rainfall in epidemic or seasonal malaria regions significantly increases the likelihood for pregnant women to be infected by malaria and that this increases the probability that their child dies before age one. Indeed, pregnant mothers who live in these regions lack immunity and therefore develop fever and anemia symptoms which lead to a higher probability of premature delivery and infant death. According to the authors, another likely mechanism is the positive correlation between malaria during pregnancy and the probability of low birth weight which is related to increased infant mortality. Ndiaye et al. (2001) also find a significant positive correlation between the total numbers of malaria deaths from September to October and the Standardised Precipitation Index (SPI) of August during the 1984-1996 period in the Niakhar HDSS where malaria is endemic². They show that the rainfall level during the rainy season (June-October) is highest in August and that the peak of malaria deaths is recorded between September and October which suggests a lagged effect of precipitation on malaria mortality which is consistent with the *Anopheles* biological breeding in the Niakhar area. Conversely, relying on DHS data of fourteen Sub-Saharan countries, Bandyopadhyay, Kanji, and Wang (2012) show that a positive deviation from the monthly rainfall average in the dry season significantly reduces the prevalence of diarrheal disease of children under age three by three percentage points. Their interpretation is that higher rainfall decreases water scarcity by making groundwater more abundant which in turn increases the availability of drinking water and clean water for households. Moreover, as higher rainfall likely increases crop yields and enhances household food security, the resulting improvement in the nutritional status of children lowers their exposure to diarrheal disease. Using two waves of Nigerian DHS data, Rabassa, Skoufias, and Jacoby (2014) show that higher rainfall has a negative contemporaneous effect on Weight-for-Height z-scores (WHZ) of children aged between 0-35 months in rural areas. More specifically, they find that a 10% positive deviation of precipitation

²Robert et al. (1998) analyzed the transmission of malaria in three villages in the Niakhar area from January to December 1995 and show that the annual biting rate of *Anopheles gambiae* which causes malaria varied from 512 to 1558 per individual depending on the villages.

from the historical mean in the year of birth rainy season lowers the WHZ by 0.034 to 0.036 standard deviations due to the high prevalence of diarrhea. Furthermore, they report that a positive rainfall shock in the rainy season before birth raises the WHZ which suggests that the positive income effect dominates the contemporaneous negative effect of water-borne-diseases. The authors stress that there is no indication of gender discrimination in the allocation of resources within households since the impact of rainfall shocks is similar for boys and girls. However, the channel linking a positive rainfall shock in the rainy season and the higher incidence of diarrhea in rural Nigeria remains unclear. Focusing on semi-arid regions of Brazil, Rocha and Soares (2015) construct a panel of municipal data for 1996-2010 combined with rainfall data to examine the impact of precipitation variability during the gestation period on birth outcomes. They find that a one standard deviation increase in precipitation in the 12 months prior to birth lowers the infant mortality rate by 1.53 points, and show that this effect is driven by a better access to safe drinkable water. According to the authors, a positive rainfall shock during the gestation period reduces water scarcity which significantly improves the health conditions of pregnant women and therefore enhances both birth outcomes and the immunological protection of the newborn. Analyzing causes of death data, they show that experiencing higher rainfall during the 12 months before birth significantly decreases deaths caused by intestinal infections and malnutrition. Furthermore, they analyze the gender heterogeneity of the effect and report that girls have a higher infant mortality rate than boys indicating that girls are more sensitive to rainfall shock during the gestational period than boys. They interpret the gender difference as a likely combination of biological and social factors. Given that parents are not able to know the gender of the child in advance of birth, i.e. during the gestation period, the authors argue that the biological factor matters more than a gender discrimination in this context. Lastly, Buchner and Rehfuess (2015) investigate the indoor air pollution channel. Using DHS data from eighteen Sub-Saharan countries, they find that the risk of suffering from Acute Lower Respiratory Infections (ALRI) for children aged under five is highest during the rainy season due to the household air pollution³. The authors show that during the rainy season, outdoor cooking is substituted for indoor cooking with coal, charcoal, and wood which raises the indoor air pollution and, as a result, the risk of ALRI for children.

To sum up, the various findings reported here are heterogeneous and appear somewhat specific to the setting. While the income-nutrition effect is likely to play a role in most rural settings, the disease-environment effect is likely to be more context dependent. Moreover, it is useful to

³A child with ALRI is defined as one who had a cough, short rapid breath, or problems in the chest or a blocked or running nose in the two weeks before the survey.

underline that the impact of income shock on mortality is indirect and operates with a time lag while the disease environment effect operates more directly and with a shorter lag. Our paper contributes to the literature by using high-quality demographic data combined with observed rainfall data to try and understand the link between rainfall and infant mortality in rural Senegal more thoroughly.

2.3 Background

The Health and Demographic Surveillance System (HDSS) of Niakhar, one of the first HDSS in Africa, was established in 1962 by the French National Research Institute for Sustainable Development (IRD) (Delaunay et al. 2013). It is located in rural Senegal in the district of Fatick, 135 km East from the capital Dakar. The size of the area is about 203 km². The population of the HDSS has been surveyed at least twice a year since 1962 and demographic events such as births, deaths, migrations have been recorded in each wave of the survey thus allowing researchers to track demographic dynamics.

The main driving factor in establishing the Niakhar HDSS was the limited availability of demographic data in Senegal after independence due to the lack of a reliable birth registration systems. Although the birth registration system has improved significantly in Senegal over the past two decades, the under-reporting of births and deaths events remains important in rural areas. According to the Demographic and Health Survey, only 50% of births in rural Senegal have been recorded in a birth registration database compared with 78% in urban areas (ANSD 2012a). The Niakhar HDSS was extended from 8 villages in 1962 and to 30 villages in 1983 located in 2 *communes* (fourth territorial division in Senegal) Ngayokheme and Diarrere. Figure 2.1 shows the study area and the location of the two rain stations. The white areas indicate flood zones that could promote mosquito breeding.

Between 1985 and 2016, the population of the HDSS has increased two-fold from 23,500 to 47,000⁴. This corresponds to an increase in the population density from 115 residents/km² in 1985 to 231 residents/km² in 2016. Niakhar's climate is typical of the Sahel with a single rainy season from June to October and a dry season from November to May. Agriculture is the main source of income for households in the district and employs about 90% of the active population (ANSD 2016). The two major crops grown in the area are millet for household food consumption and groundnut for cash. Both are highly dependent on the amount of precipitation as well as on the distribution of rainfall during the rainy season. In addition, the main concern

⁴This increase corresponds to an average population growth of 2.3% per year, very close to the national figure.

that farmers face is land salinization which affects about 33% of the area of the district (ANSD 2012b). A number of studies show that an important coping strategy to deal with the fluctuations of income due to environmental changes and smooth consumption during lean times is seasonal migration. Lalou and Delaunay (2015) analyzed the seasonal labour migration flows in the Niakhar area between 1998-2013 and show that on average 11% of men and 8% of women move out temporarily of the region each year. The authors also stress that labour migration peaks twice during the year. The first peak occurs during the dry season, in January-February, after the rain-fed harvest, when seasonal migrants move into urban areas looking for low skilled job opportunities. The second peak occurs in June, at the beginning of the rainy season, when migrants leave their homes to find work in rural areas where agricultural workers are hired.

Concerning infrastructure, the area has 4 healthcare posts located in three villages as shown in Figure 2.1 which were built in 1953, 1957, 1983, and 2014. The three oldest health infrastructures have been renovated during the last decade to improve the quality of health services. Instead rural electrification is quite recent as only 6 out of 30 villages had been electrified in 2016. Drilling and well water are the two main sources of water for household consumption. The number of public primary schools increased from 4 to 30 between 1984 and 2016.

Malaria is endemic in the area and is the leading cause of deaths for children, representing 19% of the under-five deaths (0-59 months) over the period 1985-2016. Following the Roll Back Malaria initiative, the Senegalese government launched a National Malaria Control Program in 2005 to reduce the prevalence of malaria and the associated under-five mortality. Through this program, 6 million Insecticide-Treated Nets (ITNs) were distributed between 2006 and 2010 at the national level, and more than \$130 million were spent during the period. As a result, 82% of households had at least one ITN in 2010 (WHO 2010). In the Niakhar HDSS, the massive distribution of ITNs began in June 2008. The coverage rate reached 40% in 2008 while it was less than 5% before 2008 (Trape et al. 2012). Thwing et al. (2011) report that the large-scale free distribution of ITN started in 2008 at the sub-national level and resulted in a significant increase in the coverage rate from 36.3% in 2006 to 60.4% in 2008. Furthermore, following the emergence of chloroquine resistance in malaria parasite *Plasmodium falciparum* during the 1990s (Trape et al. 1998), malaria treatment was significantly improved. Two new antimalarial drugs were introduced in the national healthcare system for the first-line treatment: Amodiaquine plus Sulfadoxine-pyrimethamine (AQ+SP) in November-December 2003 and Artesunate plus Amodiaquine (AS+AQ) in May 2006 (Trape et al. 2012). This is likely to have contributed to the decrease in under-five mortality rate in the area from 7.78 in 1985 to 0.49 in 2016.

Fertility rate remains high in the area even though the number of children per woman de-

creased slightly from 8 children per woman in 1986 to 6 in 2012 (Becker et al. 2018). According to the Demographic and Health Survey, only 10.3% of married women aged between 15 to 49 years old used contraceptives in the Fatick district in 2010 (ANSD 2012a). Although family planning programs were developed at the national level in the last decade, the prevalence of contraceptives in the Fatick district is still low, and stands at about 27.3% in 2017 (ANSD 2018).

A remarkable feature of data collected in the HDSS of Niakhar is the availability of data on causes of death through verbal autopsy. For each death, surveyors ask a set of questions to identify the circumstances and specific symptoms leading to death. Two medical doctors then analyze the questionnaire blindly and infer a probable cause of death (Etard et al. 2004; Duthé et al. 2008). In the case of discrepancy between these two doctors, a third opinion is called to solve the inconsistency or to request a discussion in order to determine the probable cause of death. This procedure ensures the reliability and the high quality of data collected in the Niakhar HDSS making it is an interesting case study to examine mortality trends as well as to try and better understand the various factors that may explain changes observed over time.

2.4 Data and Measurement of key variables

2.4.1 Mortality outcomes

The mortality data comes from the Niakhar HDSS where demographic events such as births, deaths, and migrations have been recorded at least twice a year in the 30 villages of the study area since 1983. First, a census was conducted allowing the construction of the baseline dataset which contains all HDSS residents with their individual characteristics such as sex, religion, ethnic group, marital status. Since then, the baseline dataset has been updated yearly thanks to the demographic survey conducted under the supervision of the IRD research team. All HDSS residents are followed and this longitudinal dataset allows tracking demographic dynamics over a very long time period as well as all demographic events that occurred during individual's stay in the area. Using this information, we construct a village-by-year mortality panel over the 1985-2016 period. To support the interpretation of our results, we also make use of the data on the causes of deaths presented above.

Our analysis is focused on two mortality outcomes: the Post-Neonatal Mortality Rate (PMR) and the Child Mortality Rate (CMR). We compute these two outcomes yearly at the village level over the period 1985-2016. The PMR is obtained by dividing the total number of infant deaths aged between 1 month to 11 months on the total live births at the village level. The CMR is calculated by dividing the total number of deaths between 12 and 59 months on the population

aged 12-59 months at the village level.

$$PMR_{it} = \frac{\text{Deaths infants aged 1 – 11 months}_{it}}{\text{Live Births}_{it}} \times 1000 \quad (2.1)$$

$$CMR_{it} = \frac{\text{Deaths children aged 12 – 59 months}_{it}}{\text{Population aged 12 – 59 months}_{it}} \times 1000 \quad (2.2)$$

These outcomes are computed over three different time periods to account for the strong seasonality of infant mortality. First, we calculate the PMR and CMR from June of year t to May of year $t+1$ as June corresponds to the beginning of the rainy season in Senegal. Second, we compute the PMR and CMR during the rainy season (from June to October, i.e. 5 months) and the dry season (from November to May, i.e. 7 months) as shown in Figure 2.2. To investigate gender heterogeneity, we also compute the PMR and CMR for girls and boys separately.

compared to Demographic and Health Surveys (DHS) data, the annual frequency of demographic surveys in the HDSS lowers the recall bias in demographic events allowing for high quality and precision in dating events.

2.4.2 Rainfall data

The rainfall data is provided by the *Service Départemental pour le Développement Rural (SDDR)* of the Fatick district which is a service of the Ministry of Agriculture. The data contains monthly rainfall observed over the period 1985-2016 in each *arrondissement* of the Fatick district. There is a single rainy season in Senegal that starts in June and ends in October. The 30 villages in the Niakhar HDSS are divided into the two *communes*: Ngayokheme (18 villages) and Diarrere (12 villages) which are located in the *arrondissements* of Niakhar and Tattaguine respectively. Figure 2.1 shows the location of the two *communes* and the two closest rain stations located in the Niakhar and Tattaguine *arrondissements*. The distance between the two rain stations is about 22 Km. In order to attribute rainfall at the village level, we match each village to the closest rainfall station. All 18 villages of the *commune* of Ngayokheme and 7 out of 12 villages located of the *commune* of Diarrere are matched to the Niakhar rain station. As a result, 25 villages are matched to the Niakhar rain station and 5 villages to the Tattaguine rain station.

We construct a continuous rainfall variable capturing the deviation of rainfall from the historical average in each commune. This variable is defined as follows:

$$Rain_{ct} = \ln(R_{ct}) - \ln(\bar{R}_c) \quad (2.3)$$

where R_{ct} refers to the rainfall in commune c at year t , (\bar{R}_c) is the historical mean of rainfall in commune c . $Rain_{ct}$ is the deviation between the logarithm of the yearly total rainfall and the logarithm of the historical mean of rainfall in commune c . $Rain_{ct}$ is constructed on an annual basis and can be interpreted as the percentage deviation from the historical average rainfall. This variable is widely used in the economics literature (Maccini and Yang 2009; Rabassa, Skoufias, and Jacoby 2014; Rocha and Soares 2015) to assess the impact of rainfall on health outcomes. Since we are interested in analyzing both contemporaneous and lagged effects of rainfall shocks, the rainfall data is organized so that the beginning of each year corresponds to the beginning of the rainy season. Thus, each full year is a 12-months period that starts in June of year t and ends in May of year $t+1$ as shown in Figure 2.2. Outcome variables are constructed correspondingly which is made possible by the availability of monthly data on mortality.

2.4.3 Other Data

During a research field in the Niakhar area in 2018, we collected data on infrastructures such as healthcare posts, schools, drilling stations, electrification as well as on crop yields. Health posts and drilling stations were geo-referenced and interviews provided information on their construction date. The data on schools was provided by the *Inspection d'Enseignement et de la Formation (IEF) de Fatick et Diofior* and data on electrification by the *Agence Sénégalaise d'Electrification Rurale (ASER)*. The yearly crop yield data of millet and groundnut from 1986 to 2016 at the department level of Fatick come from the *Service Départemental pour le Développement Rural (SDDR)* of Fatick.

We also compute the percentage of households who own at least one mosquito net in each village using the household goods and equipment surveys conducted in 1998, 2003, and 2014. As explained above, malaria is the main cause of under-five mortality in the Niakhar HDSS, and Insecticide Treated Nets have been shown to be an extremely effective method of malaria prevention. In 1998, only about 10% of households used a mosquito net. The coverage rate increased slightly in 2003, to about 16% before reaching 86% in 2014 thanks to the massive distribution campaign of Insecticide-Treated mosquito Nets (ITN) that started in 2008.

The daily temperature data comes from the Global Surface Summary of the Day (GSOD) provided by the US National Oceanic and Atmospheric Administration (NOAA).

2.4.4 Descriptive statistics

Table 2.1 reports summary statistics and shows the seasonal pattern of mortality in Niakhar: both post-neonatal and child mortality rates are found to be higher during the rainy season. Also, on average, at all seasons, boys have higher mortality rates than girls which is consistent with findings on infant mortality by gender (e.g. Garenne (2003)). The mean of rainfall deviation in logarithm is -0.03 indicating that on average the annual rainfall during the rainy season over the period 1985-2016 is 3% below the historical mean of rainfall. The average level of rainfall is 484 mm between 1985 and 2016, but the rainfall pattern differs across decades. The first-two decades 1985-2007 were characterized by a lower level of rainfall, 432 mm on average while the last decade 2008-2016 is characterized by a higher level of rainfall, 617 mm on average with significant inter-annual variability.

Figure 2.3 shows the evolution of post-neonatal and child mortality rates in the Niakhar HDSS over the period 1985-2016. It indicates that post-neonatal and child mortality rates have declined significantly over the period, from 81 to 7 and from 75 to 3 respectively. Looking at causes of death data, we find that measles and pertussis outbreaks drive the high mortality rates recorded in 1985. Furthermore, an outbreak of meningitis explains the peak of mortality in 1998-1999 during which PMR and CMR reached 83 and 61 in 1998, and 54 and 49 in 1999 respectively. As explained above, the introduction of two new antimalarial drugs in 2003 (AQ+SP) and 2006 (AS+AQ) together with the massive distribution of ITNs since 2008 (Trape et al. 2012) could explain the dramatic decline of both mortality outcomes observed in the 2000s. To put these statistics in perspective, we compare the under-five (0-59 months) mortality rate in the Niakhar HDSS to that of rural Senegal using Demographic and Health Survey (DHS) data. As shown in Figure 2.4, the declining mortality trend is also observed at the national level. However, while the under-five mortality rate is higher in the HDSS before 2000, the opposite pattern applies after 2000.

Figure 2.5 presents the yearly precipitations in the two *communes* over the period 1985-2016. Consistently with what has been observed at the national level (Bodian 2014), average rainfall has experienced a positive trend over the period. Variability remains high however: the Ngayokheme and Diarrere *communes* experienced the lowest amount of rainfall in 2002 (293.7 mm) and 1986 (252.8 mm) respectively and the highest in 2008 (757.8 mm) and 1995 (733.8 mm) respectively. The figure also shows that there is much more temporal variation than spatial variation between the two *communes*.

2.5 Identification strategy

Taking advantage of the exogenous variation of rainfall across communes and over time, we measure the impact of rainfall on post-neonatal and child mortality using a village-by-year panel over the period 1985-2016. Accordingly, our initial specification is the following:

$$MR_{ict} = \beta Rain_{c,t} + \alpha_i + Trend_t + \epsilon_{ict} \quad (2.4)$$

where MR_{ict} is either the Post-neonatal or the Child mortality rate in village i in commune c at time t , $Rain_{ct}$ is the log-deviation of rainfall at time t from the historical mean in commune c , α_i is a village fixed effect, and $Trend_t$ is a linear time trend.

β is the coefficient of interest which captures the effect of a deviation of rainfall from the historical mean (hereafter rainfall shock) on mortality outcomes. Following the two main channels identified in the literature review, we expect that $\beta > 0$ if the disease-environment effect dominates the income-nutrition effect, and $\beta < 0$ if the reverse is true.

The identification of an unbiased causal link between rainfall shocks and infant mortality raises a number of questions that we try to address. First, since the baseline specification includes village fixed effect, β should not be biased by any time-invariant characteristics at the village level. Second, we include a linear time trend to account for observed trends of the outcome variables of interest. Indeed, as shown in Figure 2.6, under-five mortality rates decline while rainfall exhibits a positive trend over the period 1985-2016. An alternative approach to deal with the coincidence of time trends could have been to add year fixed-effects. However, given the structure of the data, this approach would absorb all degrees of freedom, making the estimation of β impossible. Indeed, the Niakhar HDSS is a small rural area of about 203 km², and there is little spatial variability in rainfall. The problem, however, is that a linear time trend does not control for yearly shocks such as epidemic outbreaks or medical advances. In case these shocks coincide with rainfall variations, this can lead to biased estimates of the coefficient of interest β . In the remainder of this section, we discuss this concern and propose solutions to address it.

Five different types of events or shocks are likely to bias our results. First, as shown in Figure 2.4, spikes of under-five mortality were recorded in 1985 and 1998-1999. These spikes can be related to epidemic outbreaks of measles and pertussis (in 1985) and meningitis (in 1998 and 1999) that occurred in the Niakhar area. In parallel, negative rainfall shocks have been recorded in 1985, 1998, and 1999. The (possibly spurious) correlation of rainfall and mortality outcomes

that results from the coincidence of these shocks is likely to bias the coefficient of interest β . To deal with this potential bias, we include dummies capturing the epidemic outbreaks of 1985, 1998, and 1999 in our baseline specification. Second, new drugs were introduced in the Senegalese healthcare system as first-line treatments of malaria in 2003 and 2006 (Trape et al. 2012). The timing of these important medical advances in antimalarial pharmacotherapy correspond to the last decade of observation which is also characterized by higher rainfall relative to the first two decades which again could bias our results. In order to control for this coincidence, we add two year dummies for 2003 and 2006 in our baseline model. Third, the Senegalese government launched a National Malaria Control Program in 2005 which resulted in mass distribution of mosquito nets in 2008 (Thwing et al. 2011; Trape et al. 2012). At the national level, this program achieved a coverage rate of 82% in 2010 (WHO 2010). Again, the timing of this massive distribution of mosquito nets corresponds to a decade marked by increased precipitations. To control for the National Malaria Control Program, we include a dummy variable equal to one from 2008 to 2016, and zero before 2008. Fourth, the construction of infrastructures reflects the development of the Niakhar area. This is likely to be negatively associated with mortality outcomes and seems to be more frequent in the last decade characterized by higher levels of rainfall. To rule out the omitted variables issue related to development dynamic over time, we add time-varying infrastructure dummies for electrification, drilling water, and schools at the village level. We do not control for healthcare posts in the Niakhar area as three out of four were built before 1985. Finally, climate shocks such as heat waves are likely correlated with rainfall and mortality outcomes and our coefficient of interest β may capture their impact if we omit them. To rule out this concern, we include a heat wave variable indicating the number of days with a temperature above 40C in our baseline specification.

As a result, our baseline model becomes:

$$MR_{ict} = \beta Rain_{c,t} + \alpha_i + Trend_t + D_t + \gamma X_{it} + M_t + W_t + \epsilon_{ict} \quad (2.5)$$

where D_t refers to five year dummies corresponding to the three epidemic outbreaks of 1985 (measles and pertussis) and 1998-1999 (meningitis) as well as the introduction of two new antimalarial drugs in 2003 (AQ+SP) and 2006 (AS+AQ) at the national level, X_{it} is a vector of time-variant village characteristics (infrastructure dummies for electrification, drilling water, and school), M_t refers to a dummy variable controlling for the massive distribution of mosquito nets started in 2008, and W_t is the heat wave variable (the number of days with a temperature above 40C) in year t. In this specification, a causal impact of rainfall on mortality is identified

on the assumption that there are no omitted variables varying over time at the village level in the residual term ϵ_{ict} that might be correlated with both mortality outcomes and rainfall.

2.6 Results

In this section, we estimate the overall effect of rainfall on yearly post-neonatal and child mortality rates ⁵. We then measure the impact of rainfall on mortality during the rainy and dry seasons separately. Next, we analyze the impact of rainfall on mortality rates by cause of death. Lastly, we investigate the heterogeneity of effects by gender, access to health facilities, and period before presenting some robustness checks.

2.6.1 Main results

Table 2.2 presents the estimates of the impact of rainfall on mortality for all children during the 1985-2016 time period. The first column refers to mortality rates from June of year t to May of year $t+1$ (hereafter yearly rates), column (2) shows the estimates from June to October (hereafter rainy season), and column (3) presents the estimates from November to May (hereafter dry season). Panel A shows the results for the post-neonatal mortality rate (hereafter PMR) and while Panel B reports the results obtained for the child mortality rate (hereafter CMR). In the all columns, we control for village fixed effects, a linear time trend, year dummies of epidemic outbreaks, new antimalarial drugs and universal coverage with Insecticide-Treated Net (ITN), heat waves, and time-variant village-level variables - infrastructures.

Results from column 1 in Panel A indicate that a positive rainfall deviation is positively associated with PMR. Results from columns 2 and 3 suggest that the impact of rainfall on mortality is significant during the rainy season: a 10% increase in rainfall deviation from the historical mean raises the PMR during the rainy season by 2.17 points – i.e. 9.7%. Instead, the effect of rainfall deviation on PMR appears to be statistically insignificant during the dry season that follows. Turning to panel B, all coefficients are statistically insignificant which suggests that children aged 12-59 months are less sensitive to rainfall fluctuations.

To further investigate the specific pathways linking precipitation and mortality, we present evidence of the impact of precipitation on mortality by cause of death. The results obtained for PMR are presented in Table 2.3. Malaria, intestinal infections, and respiratory diseases are the three leading causes of death for post-neonatal and child mortality. Not surprisingly,

⁵Yearly rates are computed by aggregating monthly data from June of year t to May of year $t+1$ in order to synchronize yearly mortality data with the onset of the rainy season

we find that a positive rainfall deviation increases post-neonatal mortality caused by malaria and find that the impact is bigger during the rainy season. A 10% positive deviation of rainfall from the historical mean raises the PMR caused by malaria by 0.8 points – i.e. 17.9% – during the rainy season (column (2) of Panel C). Given that malaria is endemic in the Niakhar area, a positive rainfall shock during the rainy season is likely to enhance mosquito breeding in the area which drives mortality caused by malaria. Results also indicate that a 10% positive rainfall deviation increases the post-neonatal mortality caused by respiratory diseases by 1.17 points – i.e. 28.3% - during the rainy season (column (2) of Panel B). The channel through which rainfall and respiratory diseases are connected is less obvious. One possible interpretation is indoor air pollution. Indeed, Buchner and Rehfuess (2015) show that, in the sub-Saharan context, the risk of suffering from respiratory diseases is higher during the rainy season for children under five due to exposure to indoor air pollution related to cooking. According to the Niakhar household goods and equipment surveys in 2014, wood (47%) and cow dung (48%) are the two main sources of energy for cooking. It is thus plausible that higher mortality due respiratory diseases is related to the substitution of outdoor cooking to indoor cooking when rainfall is high. Another possibility is influenza as suggested by results from Niang et al. (2012). Relying on longitudinal surveillance data on influenza between 1996-2009, they show that a seasonal peak of influenza is observed among children during the rainy season. Thus, it could be that positive rainfall deviation increase post-neonatal mortality caused by respiratory diseases through the influenza channel.

Table 2.4 shows the estimates of the effect of rainfall on CMR by cause of death. Again, we find no statistically significant effect of rainfall on CMR suggesting that children aged 12-59 months are less sensitive to rainfall fluctuations.

2.6.2 Heterogeneity

2.6.2.1 Gender Heterogeneity

Understanding gender heterogeneity in child mortality is complex because it results from both biological and behavioral factors. The greater biological frailty of boys has been documented in circumstances where girls and boys have the same access to resources and care (Costa, Silva, and Victora 2017). In particular, differential vulnerability to infectious diseases between males and females has been documented in the medical literature, with males showing in general a higher susceptibility to many infectious diseases (Lunzen and Altfeld 2014). Concerning behavioral factors, they are expected to vary across regions, depending on how girls are “valued” with

respect to boys. Compared to South Asia, where numerous studies have documented higher infant mortality of girls than of boys (despite the biological disadvantage of boys), the anthropometric status of females in Sub-Saharan Africa has been found to be at par with or even better than that of males (Wamani et al. 2007). This could be related to a different division of labour and a different valuation of female and male children (Svedberg 1990). Indeed, higher female labour participation in the agricultural sector, polygamy, bride-wealth and early marriage of females are predominant customs, which results in more favourable investment in the nutritional and health status of females – or in less compensatory investment for boys. Indeed our data shows that both post-neonatal and child mortality rates are higher for boys than for girls (Table 2.5) with boys having a post-neonatal mortality risk of 42.36% vs. 33.96% for girls, and a child mortality risk of 26.33% vs. 24.07% for girls.

While boys infant mortality risk is higher, results from Panel A of Table 2.5 suggest that girls but not boys are affected by rainfall shocks. For girls under 12 months of age, a 10% positive deviation of rainfall from the historical mean raises the PMR by 2.45 points – i.e. 7.22% - over the year, and by 1.88 points – i.e. 9.17% – during the rainy season. For boys, the impact is positive but not significantly different from zero. In Panel B, the point estimates are smaller in magnitude and statistically insignificant indicating that children aged 12 to 59 months are less sensitive to rainfall shocks. In a context where infant girls have an overall lower mortality risk than boys, the interpretation of their higher sensitivity to rainfall shocks is not obvious. Looking at the structure of causes of death by gender in Table 2.6, we find that the distribution is similar concerning malaria and respiratory diseases but that intestinal infections are more often a cause of death for boys than for girls. However, regarding intestinal infections, girls have a lower probability of death compared to boys and the difference is significant at the 5% level for post-neonatal mortality. This does not provide an indication of why girls are more sensitive to weather shocks. An exploration of possible behavioral channels such as lower protection through bed nets or breastfeeding is left for future work.

2.6.2.2 Heterogeneity by Access to Health facilities

Table 2.7 reports the effects of rainfall on post-neonatal mortality by distance to health facilities.

The proximity to health facilities seem to attenuate the impact of rainfall on PMR in columns (2), (3), (5) and (6). Even if the coefficients are not significant, the magnitude of the attenuation effect is larger for villages that are closest to health infrastructures as shown in columns (2) and (5). Given the small number of villages located less than 1.5 km to health posts (3 out of 30 villages), we lack some power to detect a significant effect.

2.6.2.3 Heterogeneity by period

Table 2.8 reports the estimates of heterogeneous effects of rainfall on PMR by period. The rationale behind splitting the period in two (1985-2007 and 2008-2016) is twofold. First, malaria treatment was significantly improved in 2006 with the introduction of a new antimalarial drug in the national healthcare system Artesunate plus Amodiaquine (AS+AQ) in May 2006 (Trape et al. 2012). Second, as explained previously, the large-scale free distribution of mosquito nets started in 2008 in Senegal (Thwing et al. 2011; Trape et al. 2012).

As shown in columns (1), (2), and (3) of Table 2.8, the adverse effects of positive rainfall shocks are concentrated in the first-two decades 1985-2007. The magnitude of the effects of rainfall on PMR is larger in the first-two decades compared to the whole period 1985-2016. A 10% positive deviation of rainfall from the historical mean increases the PMR by 2.92 points – i.e. 9.87% – and 2.64 points – i.e. 9.5% – for girls in column (2) of Panels A and B respectively during the period 1985-2007 compared to 9.7% for PMR and 9.17% for girls during the whole period 1985-2016. The high prevalence of influenza in the first-two decades might explain the fact that higher rainfall increases PMR caused by respiratory diseases in the dry season. Although the seasonal peak of influenza occurred from July to September, influenza circulated during all seasons in Senegal (Niang et al. 2012). The last three columns of Table 2.8 point out interesting results. Positive rainfall shocks have no significant effects on PMR during the last decade 2008-2016 and the magnitude of effects decreases dramatically compared to the first two decades. The last decade is characterized by higher rainfall, 617 mm on average with huge inter-annual variability compared to 432 mm on average during the first two decades but the increased trend in rainfall no longer affects post-neonatal mortality. One plausible explanation is that public health policies, specifically the National Malaria Control Program performed well in the mitigation of the disease-environment effect.

Table 2.9 presents the heterogeneous effects of rainfall shocks on CMR by period. We find that positive rainfall shocks reduces the CMR caused by malaria and intestinal infections in columns (4) and (6) of panel C, suggesting that an income-nutrition effect is at play. Increased precipitations raise crop yields⁶, reduce food prices, enhance nutrient intake and household consumption, and may lead to lower child mortality caused by intestinal infections. Given that agriculture is the main source of income for households in the Niakhar area, higher rainfall is likely to increase household income allowing parents to meet health related expenses to lower the risk of malaria mortality. These findings suggest that once the disease-environment effect has been mitigated

⁶Table 2.10 shows the positive correlation between rainfall and the two major crops grown in the Fatick district: millet for household food consumption and groundnut for cash.

by public health policies (first line treatment and ITN distribution), the income-nutrition effect seems to be at work during the last decade. Although public health interventions in 2006 and 2008 have been focused on malaria, Tables 2.8 and 2.9 show a huge decrease in the average mortality rate for all causes of deaths between the first two decades and the last decade. Rossi and Villar (2020) argue that anti-malaria campaigns in Senegal have positive spillovers in malarious areas through a reallocation mechanism: households reallocate their resources to deal with other diseases. The authors show that before anti-malaria campaigns, parents spent less on preventive care (9% on consultation) and more on medication (75%) for child health. After the introduction of high subsidies in 2009, households spent more on consultation (34%) and less on medication (60%).

To sum-up, the income-nutrition effect seems to dominate the disease-environment effect during the last decade which is characterized by increased precipitations and high inter-annual variability compared to the first two decades. Public health policies appear to work and mitigate significantly the adverse effects of positive rainfall shocks on both post-neonatal and child mortality.

2.6.3 Robustness checks

2.6.3.1 Sensitivity analyses

We assess whether our results are sensitive to the exclusion of the three biggest villages of the Niakhar area (Diohine, Ngayokheme, and Toucar). These villages represent 27% of the total population in 2016 and are relatively well equipped with infrastructure. In particular, the four health facilities of the area are located in these villages. In addition, they were the first electrified villages, had access to clean water as well as school infrastructures. As shown in Table 2.11, the magnitude of the effect increases suggesting that the adverse effect of rainfall shocks is larger in smaller villages, less equipped with infrastructures. For instance, a 10% positive deviation of rainfall from the historical mean increases the PMR by 2.44 points – i.e. 10.8% – and 2.38 points – i.e. 11.6% – for girls in column (2) of Panels A and B in Table 2.11 respectively compared to 9.7% for PMR and 9.17% for girls when we include all thirty villages of the study area. Infrastructures matter in the mitigation of detrimental effects of rainfall shocks.

Furthermore, as shown in Figure 2.1, the study area contains some flood zones which could intensify mosquito breeding and the incidence of malaria. Table 2.12 shows the estimates of the impact of rainfall on PMR and CMR when we exclude thirteen villages located less than 3 km to flood zones. Although we lose some statistical power, the coefficients remain statistically

significant and increase in magnitude indicating that our results are robust when we exclude villages close to flood zones.

One possible issue is that the mortality outcomes, specifically the child mortality rate, could be driven by the out-migration flows of children aged under five. For instance, the child mortality rate could go up due to a decrease in the denominator if a negative rainfall shock (drought) is associated with higher out-migration flows of households. Fortunately, we have data on out-migration allowing us to assess the child mobility flows. Figure 2.7 shows the evolution of the out-migration rate of children aged under five years old between 1985 to 2016. The percentage of children who move out of the Niakhar area is relatively low, on average 3,80% during the period 1985-2016 indicating that out-migration flows should not be a threat for the identification.

2.6.3.2 Alternative rainfall variable

So far, the log deviation of rainfall from the historical mean is used in the econometric specification we discuss above. Alternatively, we use the level of rainfall to measure the average effect of a deviation of rainfall from the trend on mortality outcomes. We find that increased precipitations are positively associated with PMR consistent with previous results. As shown in Table 2.13, a 100 mm positive deviation of rainfall from the trend increases the PMR for all children by 4 points – i.e. 18% – and 4 points – i.e. 19% – for girls in the rainy season in column (2) of Panels A and B respectively.

2.7 Conclusion

In this paper, we investigate the impact of rainfall shocks on post-neonatal and child mortality in rural Senegal. For that purpose, we exploit the exogenous variation of rainfall over time and across communes between 1985 and 2016 and link the recorded variations with high quality demographic data collected by the health and demographic surveillance system of Niakhar.

Over the study period, we find that a positive rainfall deviation significantly increases the post-neonatal mortality rate, particularly during the rainy season. This suggests that the disease-environment effect dominates the income-nutrition effect in the region. Indeed, the analysis of data on the causes of death shows that post-neonatal mortality is largely due to malaria and respiratory diseases. Analyzing the heterogeneity of the impact by sex, we find a striking gender effect suggesting that girls, but not boys, are affected by rainfall shocks. The exploration of biological and behavioral factors that may explain this result is left to further research. We also find that the magnitude of the effect is larger in villages with less infrastructure, which could suggest

that infrastructure is important in mitigating the adverse effects of rainfall shocks on mortality. Finally, by examining the heterogeneity of the impact over time, we show that the positive impact of rainfall shocks on mortality is stronger at the beginning of the period. This is consistent with the reduction in malaria deaths following the mass distribution of insecticide-treated bed nets in the area in the mid-2000s, leading to a significant reduction in the disease-environment effect.

According to the World Malaria Report (WHO 2020), while malaria deaths decreased significantly over the last two decades, malaria remains a public health concern in many sub-Saharan countries. Furthermore, climate scientists predict that extreme rainfall events will be more frequent and rainfall variability will be higher in the future (IPCC 2018), thus raising the risk of malaria deaths in poor countries. The results in this paper provide evidence that public health interventions can be effective in mitigating the adverse effects of positive rainfall shocks despite higher precipitation recorded in the last decade. This calls for sustained efforts in the area of anti-malaria interventions in a context where these efforts could be hindered by the economic crisis associated with the COVID-19 pandemic.

REFERENCES

- Ali, Abdou and Thierry Lebel (2009). “The Sahelian standardized rainfall index revisited”. *International Journal of Climatology: A Journal of the Royal Meteorological Society* 29.12, pp. 1705–1714.
- Almond, Douglas, Janet Currie, and Valentina Duque (2018). “Childhood circumstances and adult outcomes: Act II”. *Journal of Economic Literature* 56.4, pp. 1360–1446.
- ANSD (2012a). *Enquête Démographique et de Santé à Indicateurs Multiples au Sénégal (EDS-MICS) 2010-2011*.
- (2012b). *Situation Economique et Sociale de la Région de Fatick*.
- (2016). *Situation Economique et Sociale de la Région de Fatick*.
- (2018). *Enquête Démographique et de Santé Continue au Sénégal (EDS-Continue 2017)*.
- Bandyopadhyay, Sushenjit, Shireen Kanji, and Limin Wang (2012). “The impact of rainfall and temperature variation on diarrheal prevalence in Sub-Saharan Africa”. *Applied Geography* 33, pp. 63–72.
- Becker, Charles et al. (2018). “Cinquante ans d’évolution de la fécondité à Niakhar 1963-2012. In: Delaunay V., Desclaux A., Sokhna C. (ed.), 2018. *Niakhar, mémoires et perspectives. Recherches pluridisciplinaires sur le changement en Afrique*. Marseille et Dakar, Editions de l’IRD et L’Harmattan Sénégal, 535 p.”
- Bodian, Ansoumana (2014). “Caractérisation de la variabilité temporelle récente des précipitations annuelles au Sénégal (Afrique de l’Ouest)”. *Physio-Géo. Géographie physique et environnement* Volume 8, pp. 297–312.
- Buchner, Hannes and Eva A Rehfuess (2015). “Cooking and season as risk factors for acute lower respiratory infections in African children: a cross-sectional multi-country analysis”. *PloS one* 10.6.

- Burgess, Robin et al. (2011). “Weather and death in India”. *Cambridge, United States: Massachusetts Institute of Technology, Department of Economics. Manuscript* 19.
- Costa, Janaína Calu, Inacio Crochemore Mohnsam da Silva, and Cesar Gomes Victora (2017). “Gender bias in under-five mortality in low/middle-income countries”. *BMJ global health* 2.2, e000350.
- Delaunay, Valerie et al. (2013). “Profile: the Niakhar health and demographic surveillance system”. *International journal of epidemiology* 42.4, pp. 1002–1011.
- Dell, Melissa, Benjamin F Jones, and Benjamin A Olken (2014). “What do we learn from the weather? The new climate-economy literature”. *Journal of Economic Literature* 52.3, pp. 740–98.
- Diedhiou, Arona et al. (2018). “Changes in climate extremes over West and Central Africa at 1.5 C and 2 C global warming”. *Environmental Research Letters* 13.6, p. 065020.
- Dinkelman, Taryn (2017). “Long-run Health Repercussions of Drought Shocks: Evidence from South African Homelands”. *The Economic Journal* 127.604, pp. 1906–1939.
- Duthé, Géraldine et al. (2008). “La détermination des causes de décès par autopsie verbale: étude de la mortalité palustre en zone rurale sénégalaise”. *INED, Paris, (Documents de travail no 150)* 35.
- Etard, Jean-François et al. (2004). “Childhood mortality and probable causes of death using verbal autopsy in Niakhar, Senegal, 1989–2000”. *International Journal of Epidemiology* 33.6, pp. 1286–1292.
- Flatø, Martin and Andreas Kotsadam (2014). *Droughts and gender bias in infant mortality in Sub-Saharan Africa*. Tech. rep. Memorandum.
- Garenne, Michel (2003). “Sex differences in health indicators among children in African DHS surveys”. *Journal of biosocial science* 35.4, pp. 601–614.
- Geruso, Michael and Dean Spears (2018). *Heat, Humidity, and Infant Mortality in the Developing World*. Tech. rep. National Bureau of Economic Research.
- Henry, Sabine JF and Stéphanie Dos Santos (2013). “Rainfall variations and child mortality in the Sahel: results from a comparative event history analysis in Burkina Faso and Mali”. *Population and Environment* 34.4, pp. 431–459.
- Hyland, Marie and Jason Russ (2019). “Water as destiny—The long-term impacts of drought in sub-Saharan Africa”. *World Development* 115, pp. 30–45.
- IPCC (2018). *Global warming of 1.5°C Report, Summary for Policymakers* (<http://www.ipcc.ch/report/sr15/>)
- Kim, Yeon Soo (2010). “The impact of rainfall on early child health”. *Unpublished manuscript*.

- Kudamatsu, Masayuki, Torsten Persson, and David Strömberg (2012). “Weather and infant mortality in Africa”. *CEPR Discussion Paper No. DP9222*.
- Kumar, Santosh, Ramona Molitor, and Sebastian Vollmer (2016). “Drought and early child health in rural India”. *Population and Development Review*, pp. 53–68.
- Lalou, Richard and Valérie Delaunay (2015). “Migrations saisonnières et changement climatique en milieu rural sénégalais: forme ou échec de l’adaptation? In: Sultan Benjamin (ed.), Lalou Richard (ed.), Amadou Sanni M. (ed.), Oumarou A. (ed.), Soumaré M.A. (ed.) *Les sociétés rurales face aux changements climatiques et environnementaux en Afrique de l’Ouest*. Marseille : IRD, 287-313. (Synthèses). ISBN 978-2-7099-2146-6”.
- Lunzen, Jan van and Marcus Altfeld (2014). “Sex differences in infectious diseases—common but neglected”. *The Journal of infectious diseases* 209.suppl_3, S79–S80.
- Maccini, Sharon and Dean Yang (2009). “Under the weather: Health, schooling, and economic consequences of early-life rainfall”. *American Economic Review* 99.3, pp. 1006–26.
- Ndiaye, Ousmane et al. (2001). “Variations climatiques et mortalité attribuée au paludisme dans la zone de Niakhar, Sénégal, de 1984 à 1996”. *Cahiers d’études et de recherches francophones/Santé* 11.1, pp. 25–33.
- Niang, Mbayame Ndiaye et al. (2012). “Sentinel surveillance for influenza in Senegal, 1996–2009”. *The Journal of infectious diseases* 206.suppl_1, S129–S135.
- Rabassa, Mariano, Emmanuel Skoufias, and Hanan Jacoby (Apr. 2014). “Weather and Child Health in Rural Nigeria”. *Journal of African Economies* 23.4, pp. 464–492. ISSN: 0963-8024. DOI: 10.1093/jae/eju005. eprint: <https://academic.oup.com/jae/article-pdf/23/4/464/2241888/eju005.pdf>.
- Robert, Vincent et al. (1998). “La transmission du paludisme dans la zone de Niakhar, Sénégal”. *Tropical Medicine & International Health* 3.8, pp. 667–677.
- Rocha, Rudi and Rodrigo R Soares (2015). “Water scarcity and birth outcomes in the Brazilian semiarid”. *Journal of Development Economics* 112, pp. 72–91.
- Rossi, Pauline and Paola Villar (2020). “Private health investments under competing risks: evidence from malaria control in Senegal”. *Journal of Health Economics* 73, p. 102330.
- Shah, Manisha and Bryce Millett Steinberg (2017). “Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital”. *Journal of Political Economy* 125.2, pp. 527–561.
- Svedberg, Peter (1990). “Undernutrition in Sub-Saharan Africa: Is there a gender bias?” *The Journal of Development Studies* 26.3, pp. 469–486.

- Sylla, Mouhamadou Bamba et al. (2015). “Projected changes in the annual cycle of high-intensity precipitation events over West Africa for the late twenty-first century”. *Journal of Climate* 28.16, pp. 6475–6488.
- Thwing, Julie I et al. (2011). “Success of Senegal’s first nationwide distribution of long-lasting insecticide-treated nets to children under five-contribution toward universal coverage”. *Malaria Journal* 10.1, pp. 1–8.
- Trape, Jean-François et al. (1998). “Impact of chloroquine resistance on malaria mortality”. *Comptes Rendus de l’Académie des Sciences-Series III-Sciences de la Vie* 321.8, pp. 689–697.
- Trape, Jean-François et al. (2012). “New malaria-control policies and child mortality in senegal: reaching millennium development goal 4”. *Journal of Infectious Diseases* 205.4, pp. 672–679.
- UNISDR (2015). *Weather Disasters Report: The human cost of weather related disasters 1995-2015*.
- Wamani, Henry et al. (2007). “Boys are more stunted than girls in sub-Saharan Africa: a meta-analysis of 16 demographic and health surveys”. *BMC pediatrics* 7.1, pp. 1–10.
- WHO (2010). *Roll Back Malaria, Focus on Senegal. Progress Impact Series, number 4*.
- (2015). *World Health Statistics 2015*.
- (2020). *World Malaria Report 2020, 20 years of global progress and challenges*.

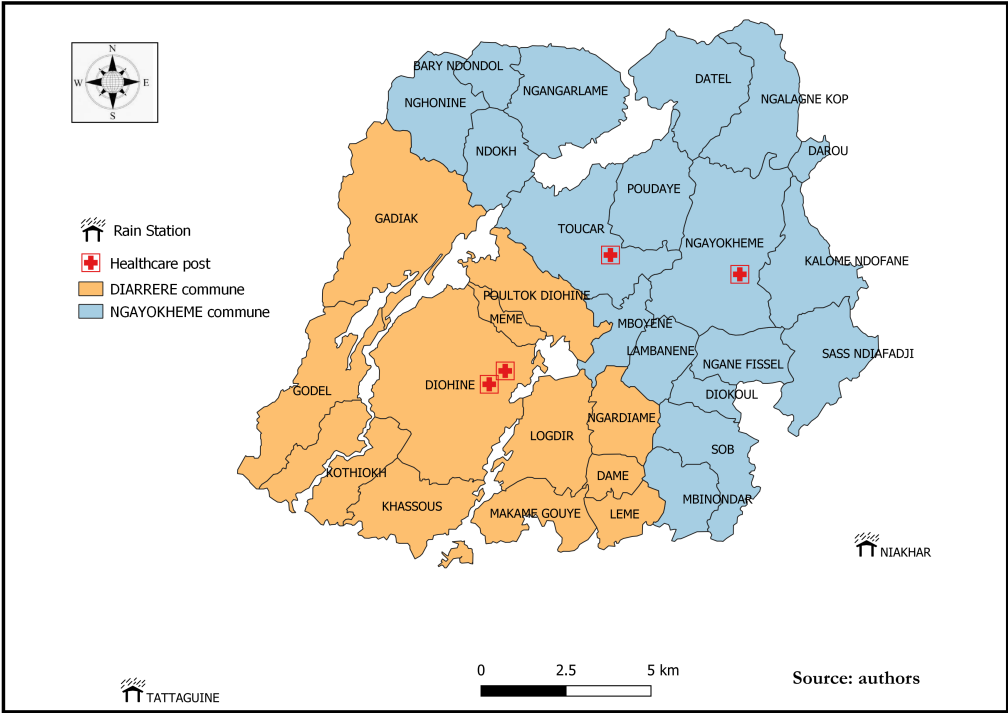


Figure 2.1: Location of Niakhar area and Rain station.

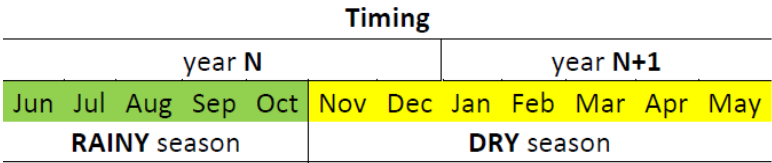


Figure 2.2: Timing of outcomes

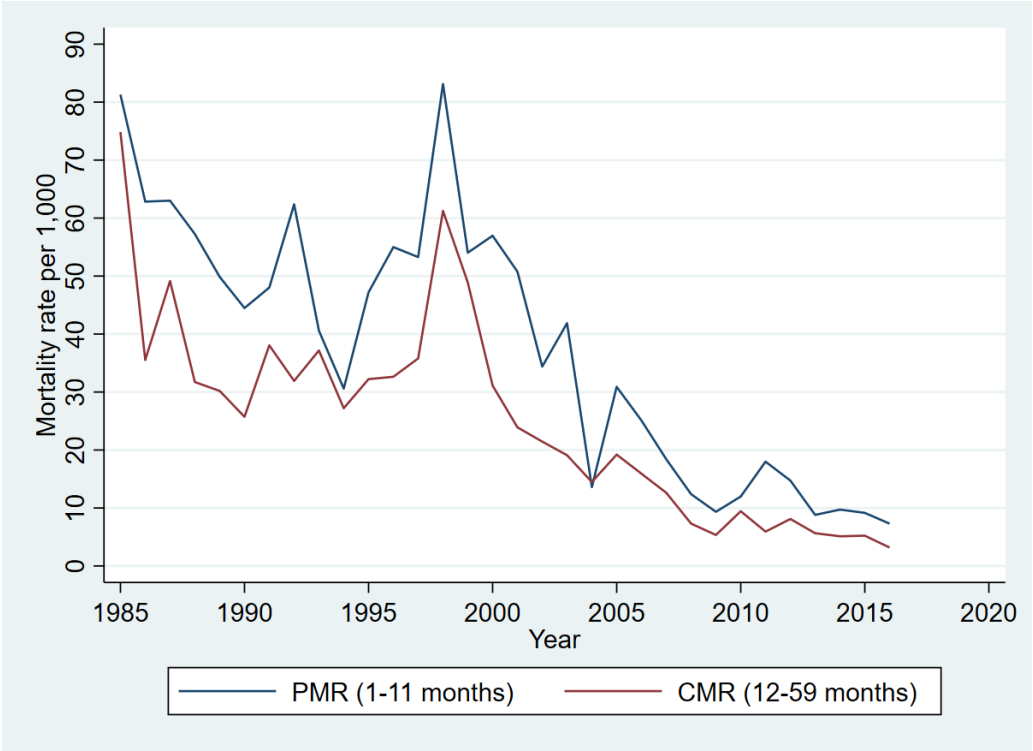


Figure 2.3: Evolution of Post-Neonatal and Child Mortality rates in the Niakhar HDSS 1985-2016.

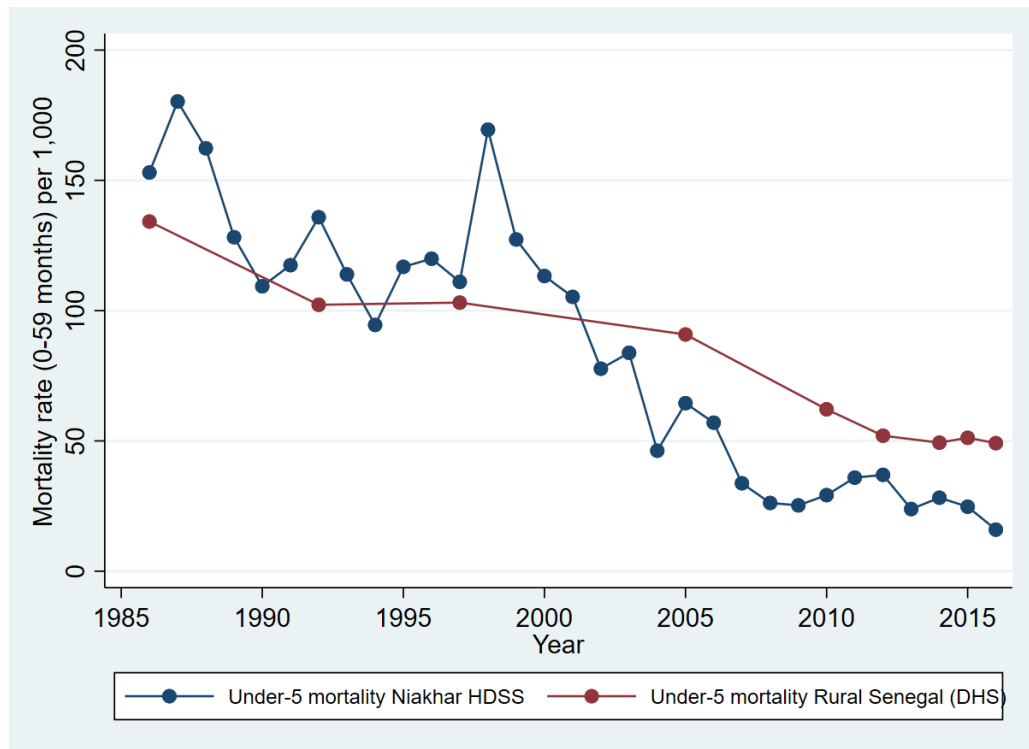


Figure 2.4: DHS data compared to the Niakhar HDSS data.

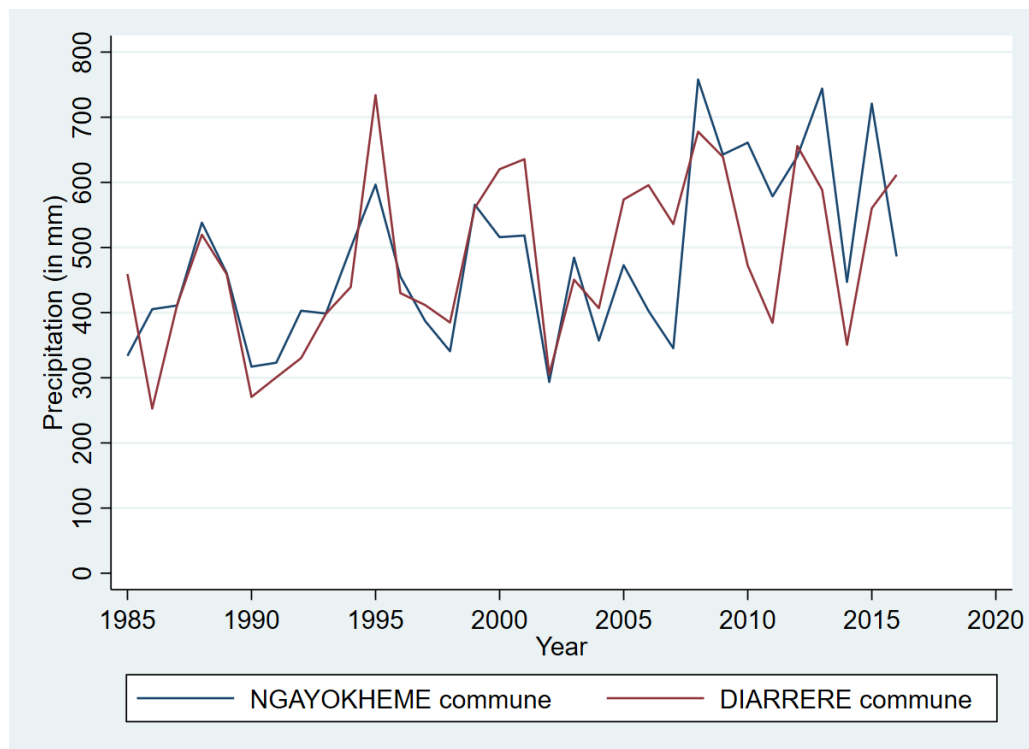


Figure 2.5: Yearly precipitation (in mm) in the 2 communes of Niakhar area.

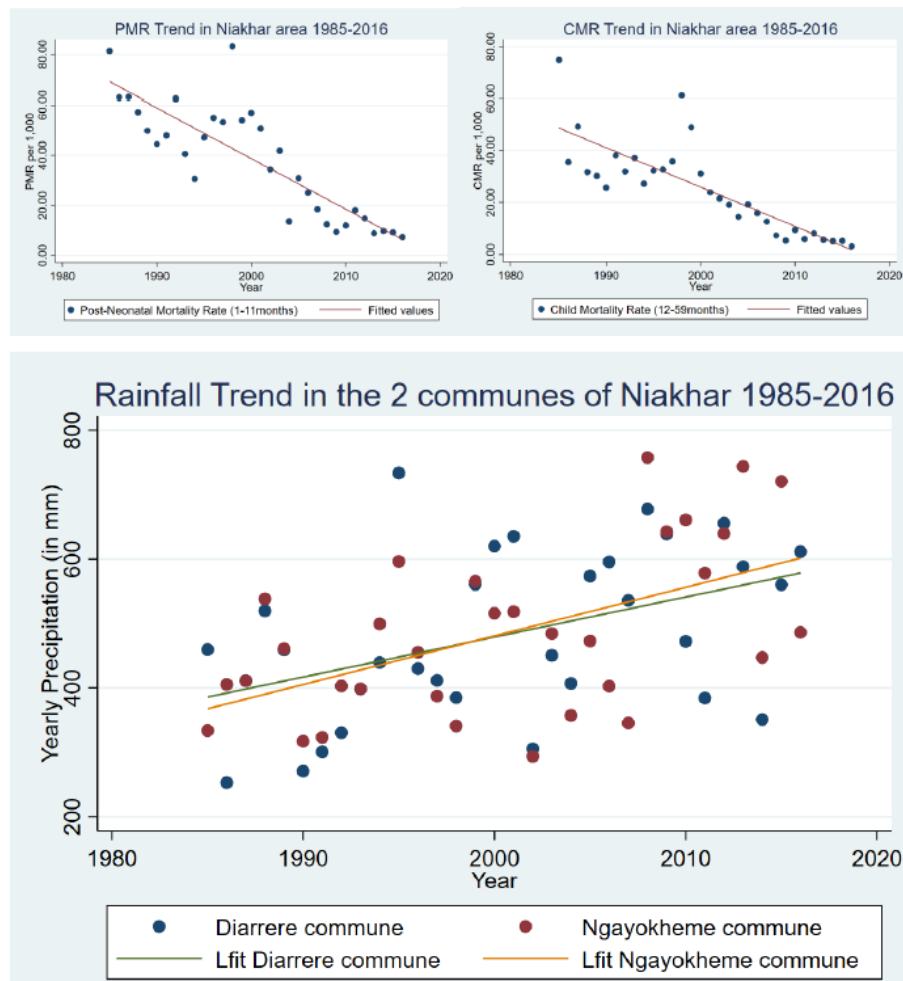


Figure 2.6: Trend of mortality outcomes and rainfall over time 1985-2016 in the Niakhar area.

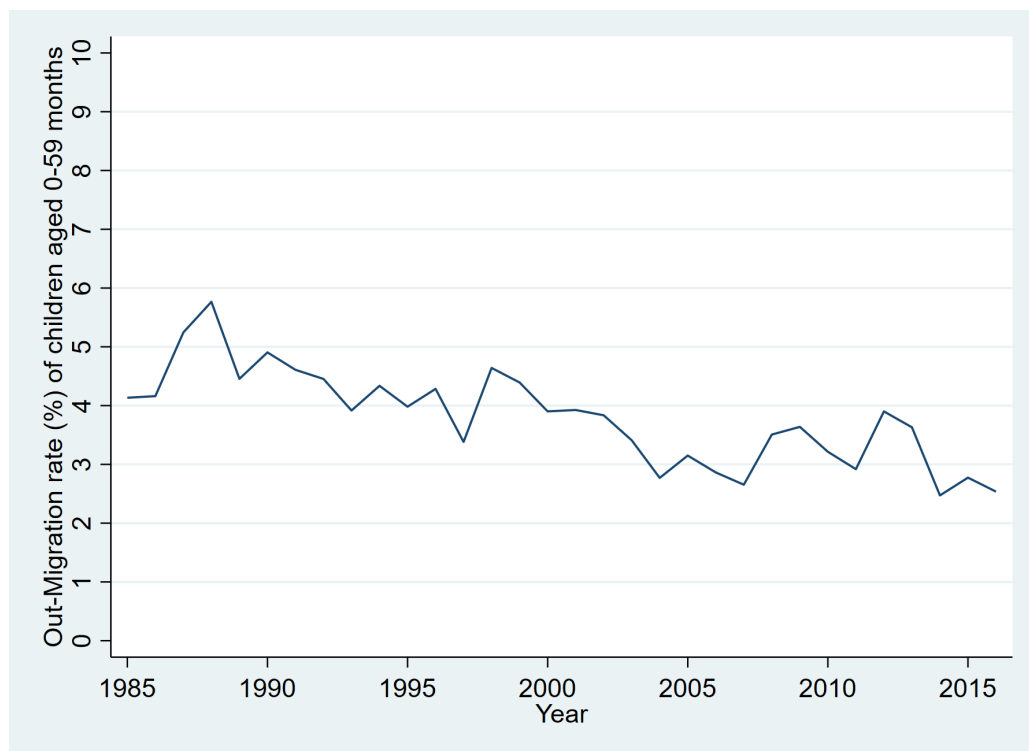


Figure 2.7: Out-Migration rate of children under 5 per year between 1985-2016.

Table 2.1: Descriptive statistics

	Mean	Std. Deviation	Min	Max	N
Post-Neonatal Mortality rate (1-11 months)					
June-May	38.33	(57.75)	0	1000	960
June-Oct (Rainy season)	22.31	(49.66)	0	1000	960
Nov-May (Dry season)	17.65	(44.21)	0	1000	960
Child Mortality rate (12-59 months)					
June-May	25.53	(30.52)	0	500	960
June-Oct (Rainy season)	15.59	(21.35)	0	333	960
Nov-May (Dry season)	10.01	(15.44)	0	200	960
Rainfall					
Level of Rainfall (in mm)	484.11	(128.01)	253	758	960
Log Rain deviation	-0.03	(0.26)	-0.65	0.45	960
Temperature					
Number of days with a Temperature > 40C per year	53.31	(14.23)	25	83	960
Village characteristics					
Electrification	0.05	(0.21)	0	1	960
Drilling water	0.18	(0.38)	0	1	960
School	0.51	(0.50)	0	1	960

Table 2.2: Impact of rainfall on Post-Neonatal and Child Mortality

	Full Year (1)	Rainy season (2)	Dry season (3)
Panel A: PMR (1-11 months)			
Log rain deviation	19.52** (8.91)	21.76*** (7.98)	3.33 (7.18)
Mean PMR	38.33	22.31	17.65
Panel B: CMR (12-59 months)			
Log rain deviation	3.12 (4.14)	4.16 (3.11)	-1.37 (2.26)
Mean CMR	25.53	15.59	10.01
N	960	960	960

Notes: The dependent variable is Post-Neonatal mortality rate per 1,000 in Panel A, and Child mortality rate per 1,000 in Panel B. All regressions include Village fixed-effects, Linear time trend, Epidemic outbreak dummies, New antimalarial drug dummy, Heat waves (number of days with a temperature above 40C), Insecticide-Treated Net (ITN) dummy, Village level controls: time-varying infrastructure dummies (school, electrification, drilling water). Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 2.3: Impact of rainfall on Post-Neonatal Mortality by Cause of death

	Full Year (1)	Rainy season (2)	Dry season (3)
Panel A: Malaria			
Log rain deviation	9.65*** (3.18)	8.00** (3.58)	2.78* (1.52)
Mean PMR	6.13	4.46	1.91
Panel B: Respiratory diseases			
Log rain deviation	8.44 (5.23)	11.73** (5.88)	2.53 (1.96)
Mean PMR	7.59	4.14	3.20
Panel C: Intestinal infections			
Log rain deviation	-0.22 (4.59)	-0.89 (3.28)	0.68 (1.72)
Mean PMR	9.80	6.51	3.18
N	960	960	960

Notes: The dependent variable is Post-Neonatal mortality rate per 1,000 in Panel A, and Child mortality rate per 1,000 in Panel B. All regressions include Village fixed-effects, Linear time trend, Epidemic outbreak dummies, New antimalarial drug dummy, Heat waves (number of days with a temperature above 40C), Insecticide-Treated Net (ITN) dummy, Village level controls: time-varying infrastructure dummies (school, electrification, drilling water). Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 2.4: Impact of rainfall on Child Mortality by Cause of death

	Full Year (1)	Rainy season (2)	Dry season (3)
Panel A: Malaria			
Log rain deviation	3.16 (2.47)	2.53 (2.26)	0.47 (0.69)
Mean CMR	6.81	5.01	1.77
Panel B: Respiratory diseases			
Log rain deviation	1.43 (0.97)	0.57 (0.51)	0.85 (0.90)
Mean CMR	1.80	0.74	1.06
Panel C: Intestinal infections			
Log rain deviation	-2.36 (2.15)	-0.09 (1.77)	-1.91 (1.29)
Mean CMR	6.97	4.55	2.24
N	960	960	960

Notes: The dependent variable is Post-Neonatal mortality rate per 1,000 in Panel A, and Child mortality rate per 1,000 in Panel B. All regressions include Village fixed-effects, Linear time trend, Epidemic outbreak dummies, New antimalarial drug dummy, Heat waves (number of days with a temperature above 40C), Insecticide-Treated Net (ITN) dummy, Village level controls: time-varying infrastructure dummies (school, electrification, drilling water). Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 2.5: Impact of rainfall on Post-Neonatal and Child Mortality by Gender

	Full Year (1)	Rainy season (2)	Dry season (3)
Panel A: PMR (1-11 months)			
Girl			
Log rain deviation	24.53** (9.69)	18.80** (8.58)	6.56 (5.65)
Mean PMR Girl	33.96	20.50	14.18
Boy			
Log rain deviation	18.98 (13.05)	8.46 (7.56)	2.41 (9.38)
Mean PMR Boy	42.36	21.66	20.19
Panel B: CMR (12-59 months)			
Girl			
Log rain deviation	3.18 (4.29)	4.54 (3.22)	-1.74 (2.74)
Mean CMR Girl	24.07	14.73	9.35
Boy			
Log rain deviation	2.36 (5.80)	3.11 (4.36)	-0.97 (3.21)
Mean CMR Boy	26.33	15.78	10.57
N	960	960	960

Notes: The dependent variable is Post-Neonatal mortality rate per 1,000 in Panel A, and Child mortality rate per 1,000 in Panel B. All regressions include Village fixed-effects, Linear time trend, Epidemic outbreak dummies, New antimalarial drug dummy, Heat waves (number of days with a temperature above 40C), Insecticide-Treated Net (ITN) dummy, Village level controls: time-varying infrastructure dummies (school, electrification, drilling water). Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 2.6: Descriptive statistics on post-neonatal and child mortality by cause of death and gender

Panel A: Post-neonatal mortality	Boy N=760	Girl N=675	Diff.	Std. Error	N
Malaria	0.15	0.16	-0.01	0.02	1435
Respiratory diseases	0.16	0.18	-0.02	0.02	1435
Intestinal infections	0.29	0.23	0.06**	0.02	1435
Panel B: Child mortality	Boy N=1769	Girl N=1601	Diff.	Std. Error	N
Malaria	0.26	0.28	-0.02	0.02	3370
Respiratory diseases	0.06	0.07	-0.01	0.01	3370
Intestinal infections	0.27	0.25	0.02	0.02	3370

Notes: Panel A shows the probability of post-neonatal mortality (1-11 months) by cause of death and gender. Panel B reports the probability of child mortality (12-59 months) by cause of death and gender.

Table 2.7: Heterogeneous effects of rainfall on Post-Neonatal and Child Mortality by distance to Health facilities

	Full Year			Rainy season		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: PMR (1-11 months)						
Log rain deviation	14.38 (18.18)	22.23** (9.33)	21.46** (10.05)	22.30 (16.28)	23.58*** (8.35)	22.68** (8.99)
Log rain deviation*Distance (km) to the Health post	1.32 (4.08)			-0.14 (3.65)		
Log rain deviation*Less than 1.5km to the Health post		-23.12 (23.56)			-15.52 (21.10)	
Log rain deviation*Less than 2.5km to the Health post			-6.01 (14.29)			-2.86 (12.79)
Mean PMR	38.33			22.31		
Panel B: CMR (12-59 months)						
Log rain deviation	3.15 (8.45)	3.99 (4.34)	4.10 (4.67)	6.31 (6.34)	4.54 (3.25)	3.47 (3.50)
Log rain deviation*Distance (km) to the Health post	-0.01 (1.90)			-0.55 (1.42)		
Log rain deviation*Less than 1.5km to the Health post		-7.40 (10.96)			-3.27 (8.21)	
Log rain deviation*Less than 2.5km to the Health post			-3.01 (6.64)			2.12 (4.98)
Mean CMR	25.53			15.59		
N	960	960	960	960	960	960

Notes: The dependent variable is Post-Neonatal mortality rate per 1,000. All regressions include Village fixed-effects, Linear time trend, Epidemic outbreak dummies, New antimalarial drug dummy, Heat waves (number of days with a temperature above 40C), Insecticide-Treated Net (ITN) dummy, Village level controls: time-varying infrastructure dummies (school, electrification, drilling water). Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 2.8: Heterogeneous effects of rainfall on Post-Neonatal Mortality by period

	Period 1985–2007			Last decade 2008–2016		
	Full Year (1)	Rainy season (2)	Dry season (3)	Full Year (4)	Rainy season (5)	Dry season (6)
Panel A: PMR (1-11 months) - All						
Log rain deviation	31.25*** (12.10)	29.26*** (10.95)	11.36 (9.81)	1.95 (8.59)	3.24 (6.34)	-1.43 (6.94)
Mean PMR	50.24	29.64	22.88	12.13	6.16	6.14
Panel B: PMR by Gender						
Girl						
Log rain deviation	39.15*** (13.37)	26.44** (11.94)	14.83* (7.76)	-7.82 (6.36)	1.66 (4.64)	-10.23** (4.66)
Mean PMR Girl	44.97	27.82	18.15	9.74	4.39	5.43
Boy						
Log rain deviation	26.93 (17.47)	8.03 (9.94)	8.79 (12.47)	9.65 (15.95)	4.91 (10.78)	5.92 (13.13)
Mean PMR Boy	55.19	27.94	26.42	14.13	7.83	6.48
Panel C: PMR by Cause of Death						
Malaria						
Log rain deviation	12.53*** (4.26)	9.86** (4.80)	4.20** (2.11)	0.99 (3.78)	1.16 (4.35)	0.02 (0.59)
Mean Malaria	8.27	5.90	2.70	1.45	1.28	0.16
Respiratory diseases						
Log rain deviation	13.99* (7.27)	16.40** (8.26)	4.82* (2.70)	-2.05 (2.22)	0.50 (0.77)	-1.96 (1.45)
Mean Respiratory diseases	10.58	5.87	4.43	1.01	0.30	0.47
Intestinal infections						
Log rain deviation	2.95 (6.31)	-0.06 (4.46)	2.59 (2.33)	-1.17 (3.80)	-1.20 (3.55)	-0.39 (1.71)
Mean Intestinal infections	13.28	8.84	4.26	2.14	1.37	0.79
N	690	690	690	270	270	270

Notes: The dependent variable is Post-Neonatal mortality rate per 1,000. All regressions include Village fixed-effects, Linear time trend, Epidemic outbreak dummies, New antimalarial drug dummy, Heat waves (number of days with a temperature above 40C), Village level controls: time-varying infrastructure dummies (school, electrification, drilling water). Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 2.9: Heterogeneous effects of rainfall on Child Mortality by period

	Period 1985–2007			Last decade 2008–2016		
	Full Year (1)	Rainy season (2)	Dry season (3)	Full Year (4)	Rainy season (5)	Dry season (6)
Panel A: CMR (12-59 months) - All						
Log rain deviation	3.11 (5.67)	4.92 (4.28)	-2.04 (3.05)	0.72 (3.65)	0.18 (2.39)	0.32 (2.64)
Mean CMR	34.12	20.86	13.35	6.61	3.97	2.65
Panel B: CMR by Gender						
Girl						
Log rain deviation	5.03 (5.81)	6.45 (4.40)	-1.95 (3.68)	-0.11 (4.46)	0.28 (3.18)	-0.77 (3.27)
Mean CMR Girl	32.12	19.63	12.48	6.34	3.93	2.44
Boy						
Log rain deviation	0.22 (7.92)	2.50 (5.98)	-2.31 (4.34)	2.86 (5.10)	1.02 (3.53)	1.69 (3.86)
Mean CMR Boy	35.16	21.10	14.11	6.91	4.07	2.77
Panel C: CMR by Cause of Death						
Malaria						
Log rain deviation	3.31 (3.44)	2.93 (3.13)	0.37 (0.96)	-2.41* (1.30)	-2.14 (1.30)	-0.55* (0.33)
Mean Malaria	9.30	6.80	2.45	1.32	1.05	0.26
Respiratory diseases						
Log rain deviation	2.15 (1.35)	0.80 (0.71)	1.46 (1.27)	0.28 (0.52)	-0.06 (0.29)	-0.05 (0.20)
Mean Respiratory diseases	2.51	1.01	1.51	0.24	0.12	0.06
Intestinal infections						
Log rain deviation	-2.92 (2.96)	-0.23 (2.47)	-2.15 (1.81)	-0.47 (1.83)	0.30 (0.98)	-0.80* (0.46)
Mean Intestinal infections	9.56	6.29	3.09	1.27	0.70	0.34
N	690	690	690	270	270	270

Notes: The dependent variable is Child mortality rate per 1,000. All regressions include Village fixed-effects, Linear time trend, Epidemic outbreak dummies, New antimalarial drug dummy, Heat waves (number of days with a temperature above 40C), Village level controls: time-varying infrastructure dummies (school, electrification, drilling water). Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 2.10: Correlation between Rainfall and Crop yields in the department of Fatick 1986-2016

	Millet (1)	Groundnut (2)
Log rain deviation	382.56*** (118.43)	488.10*** (163.09)
Mean crop yield (kg/ha)	655.88	747.34
N	31	31
R ²	0.26	0.24

Notes: Dependent variables are Millet's crop yields (kg/ha) and Groundnut's crop yields (kg/ha) measured at the department level of Fatick (only temporal variation 1986-2016). Independent variable is the logarithm of rainfall deviation from the historical mean at the department level. Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 2.11: Impact of rainfall on Post-Neonatal and Child Mortality with sample restriction: excluding the 3 biggest villages, 1985-2016

	PMR (1-11 months)			CMR (12-59 months)		
	Full Year (1)	Rainy season (2)	Dry season (3)	Full Year (4)	Rainy season (5)	Dry season (6)
Panel A: All						
Log rain deviation	23.52** (9.95)	24.43*** (8.94)	5.52 (8.03)	3.56 (4.61)	4.35 (3.46)	-1.17 (2.50)
Mean	38.70	22.60	17.90	25.49	15.57	9.97
Panel B: Gender						
Girl						
Log rain deviation	31.90*** (10.77)	23.86** (9.58)	9.25 (6.27)	3.10 (4.77)	4.48 (3.58)	-1.85 (3.04)
Mean Girl	33.64	20.51	13.93	24.02	14.73	9.28
Boy						
Log rain deviation	20.28 (14.60)	6.65 (8.43)	4.55 (10.46)	3.21 (6.44)	3.46 (4.86)	-0.46 (3.56)
Mean Boy	43.36	21.92	20.85	26.22	15.67	10.53
Panel C: Cause of Death						
Malaria						
Log rain deviation	10.85*** (3.55)	8.58** (4.01)	3.57** (1.69)	3.37 (2.76)	2.55 (2.53)	0.63 (0.76)
Mean Malaria	6.25	4.52	1.98	6.85	5.09	1.72
Respiratory diseases						
Log rain deviation	9.39 (5.86)	12.86* (6.61)	3.16 (2.16)	1.60 (1.08)	0.61 (0.57)	0.97 (1.01)
Mean Respiratory diseases	7.73	4.30	3.15	1.83	0.75	1.08
Intestinal infections						
Log rain deviation	-0.06 (5.12)	-1.05 (3.67)	1.04 (1.90)	-2.61 (2.39)	-0.17 (1.98)	-2.07 (1.44)
Mean Intestinal infections	9.75	6.53	3.10	6.95	4.49	2.23
N	864	864	864	864	864	864

Notes: The dependent variable is Post-Neonatal mortality rate per 1,000. All regressions include Village fixed-effects, Linear time trend, Epidemic outbreak dummies, New antimalarial drug dummy, Heat waves (number of days with a temperature above 40C), Insecticide-Treated Net (ITN) dummy, Village level controls: time-varying infrastructure dummies (school, electrification, drilling water). Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 2.12: Impact of rainfall on Post-Neonatal and Child Mortality with sample restriction: excluding 13 villages close to flood zones, 1985-2016

	PMR (1-11 months)			CMR (12-59 months)		
	Full Year	Rainy season	Dry season	Full Year	Rainy season	Dry season
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All						
Log rain deviation	26.49*	30.55**	4.97	2.73	1.56	0.49
	(15.04)	(13.77)	(12.36)	(6.78)	(5.13)	(3.56)
Mean	37.88	22.85	17.54	25.41	15.66	9.81
Panel B: Gender						
Girl						
Log rain deviation	39.55**	31.79**	10.30	0.23	-0.45	-0.17
	(16.36)	(14.98)	(8.63)	(6.41)	(4.75)	(4.27)
Mean Girl	34.86	22.25	13.63	23.82	14.45	9.41
Boy						
Log rain deviation	17.87	11.45	6.24	4.34	2.38	1.41
	(17.01)	(10.48)	(14.22)	(9.60)	(7.29)	(4.97)
Mean Boy	38.02	19.43	19.26	25.92	15.91	10.01
Panel C: Cause of Death						
Malaria						
Log rain deviation	14.75***	13.87**	3.77	1.68	0.68	0.69
	(5.44)	(6.39)	(2.50)	(4.30)	(3.97)	(1.02)
Mean Malaria	6.31	4.76	1.94	6.71	5.05	1.60
Respiratory diseases						
Log rain deviation	7.16	16.16	0.84	2.72	0.70	1.89
	(9.04)	(10.47)	(2.65)	(1.66)	(0.84)	(1.62)
Mean Respiratory diseases	8.03	4.30	3.26	1.85	0.76	1.09
Intestinal infections						
Log rain deviation	-2.24	-3.30	1.56	-2.46	-0.18	-2.09
	(8.03)	(5.63)	(2.36)	(3.64)	(3.04)	(2.29)
Mean Intestinal infections	9.67	6.88	2.51	7.76	4.94	2.55
N	544	544	544	544	544	544

Notes: The dependent variable is Post-Neonatal mortality rate per 1,000. All regressions include Village fixed-effects, Linear time trend, Epidemic outbreak dummies, New antimalarial drug dummy, Heat waves (number of days with a temperature above 40C), Insecticide-Treated Net (ITN) dummy, Village level controls: time-varying infrastructure dummies (school, electrification, drilling water). Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 2.13: Impact of rainfall on Post-Neonatal and Child Mortality using the level of rainfall, 1985-2016

	PMR (1-11 months)			CMR (12-59 months)		
	Full Year	Rainy season	Dry season	Full Year	Rainy season	Dry season
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All						
Rainfall	0.03*	0.04**	0.00	0.01	0.01	-0.00
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.00)
Mean	38.33	22.31	17.65	25.53	15.59	10.01
Panel B: Gender						
Girl						
Rainfall	0.04**	0.04**	0.01	0.01	0.01	-0.00
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Mean Girl	33.96	20.50	14.18	24.07	14.73	9.35
Boy						
Rainfall	0.04	0.02	0.00	0.00	0.01	-0.00
	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Mean Boy	42.36	21.66	20.19	26.33	15.78	10.57
Panel C: Cause of Death						
Malaria						
Rainfall	0.02***	0.02**	0.01	0.01	0.00	0.00
	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
Mean Malaria	6.13	4.46	1.91	6.81	5.01	1.77
Respiratory diseases						
Rainfall	0.01	0.02*	0.00	0.00	0.00	0.00
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Mean Respiratory diseases	7.59	4.14	3.20	1.80	0.74	1.06
Intestinal infections						
Rainfall	-0.00	-0.00	0.00	-0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Mean Intestinal infections	9.80	6.51	3.18	6.97	4.55	2.24
N	960	960	960	960	960	960

Notes: Dependent variables are Post-Neonatal mortality rate per 1,000, and Child mortality rate per 1,000. All regressions include Village fixed-effects, Linear time trend, Epidemic outbreak dummies, New antimalarial drug dummy, Heat waves (number of days with a temperature above 40C), Insecticide-Treated Net (ITN) dummy, Village level controls: time-varying infrastructure dummies (school, electrification, drilling water). Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

CHAPTER 3

CHILDREN'S MIGRATION AND SCHOOLING IN SOUTH AFRICA

Abstract

More than one out of three children aged under fifteen years old moved within the country between 2008 and 2015 in South Africa. This paper investigates the effect of children's migration, especially moving with your biological mother on educational outcomes in South Africa. Using the five waves of the National Income Dynamics Study (NIDS) individual panel data, I compare the school progression and educational attainment of migrant and non-migrant siblings. Controlling for individual unobserved heterogeneity allows us to deal with part of the endogeneity issue of migration. Results show that children who moved at early ages with their mothers have a better school progression and are more likely to achieve grade 12.

Keywords: Migration, Schooling, South Africa.

3.1 Introduction

Children are highly mobile in South Africa and their mobility is strongly associated with the mother's labour migration (Madhavan et al. 2012; Hall and Posel 2019; Hall and Posel 2020). About 2.5 million black children aged under 15 migrated within the country between 2008 and 2015 corresponding to more than one out of three children, and half of children's moves ended-up in the co-residence with their mothers (Hall and Posel 2020). The migration patterns of children are complex and vary from co-migration with their moms, moving to join their mothers, or moving away from their mothers. However, little is known about the trajectory in terms of schooling and educational attainment of migrant children, in particular those who co-migrate with their moms or move to join their mothers compared to the one of non-migrant children. This paper investigates to what extent the children's migration within South Africa may affect their educational outcomes.

Relying on the first four waves of the NIDS individual panel data, Hall and Posel (2019) examine the relationship between migration patterns of children aged under fifteen years old and maternal mobility in South Africa. The authors find that children with a migrant mother are more likely to migrate than those with a non-migrant mother and most of the child and maternal mobility occurs within urban or rural areas. Furthermore, they provide evidence suggesting that children whose mothers are actively looking for a job are two-and-a-half times more likely to move than children whose mothers are unemployed and do not look for work. Investigating the drivers of child migration, the authors argue that once a mother finds a secure job and a housing arrangement in the destination area, especially in an urban area, the children move to join their mother to have access to better schooling opportunities. Hall and Posel (2020) find that 75% of children aged under fifteen years old who migrate to join their mothers move to urban areas. Most of the existing evidence in the economic literature focus on the effect of parental migration on children left behind (Amuedo-Dorantes and Pozo 2010; Antman 2011; Antman 2012; Zhang et al. 2014; Meng and Yamauchi 2017). One common hypothesis of these studies is that migrant remittances allow the household of origin to fund the schooling fees of children left behind. They also stress that parental absence adversely affects the child's cognitive development and may reduce the time allocated to school and raise work hours to deal with financial hardship in the short run in the household. Furthermore, maternal labour migration might lead to child fostering to extended family members to be cared or for educational opportunities resulting in the child's move in a host family. Using household-level data from the South African Project for Statistics on Living Standards and Development, Zimmerman (2003) shows that human

capital investment is one of the main motives of black child fostering. The author finds that the migration of a child from the biological family to a host family reduces the risk of not being enrolled in school by 25 percent compared to the one if the child had stayed in their biological household.

Another strand of the literature examines the trajectory of children who move with their family in the context of high income countries. Chyn and Katz (2021) review the existing evidence on the impact of children's move to lower-poverty neighborhoods on long-run educational outcomes in developed countries. They show that children who migrated with their families to a better neighborhood environment in the United States, especially at early ages, are more likely to attend college and have higher educational attainment in later life. School quality through increased school resources and better-performing schools in the destination area seems to be the pathways at work. Chetty, Hendren, and Katz (2016) assess the causal effect of the Moving to Opportunity experiment between 1994 and 1998 of the US Department of Housing on children's long-term economic, schooling, and demographic outcomes. The experiment aimed to offer randomly selected families "subsidized housing vouchers" to migrate from high-poverty areas to lower-poverty neighborhoods. Using the Moving to Opportunity data combined with the federal income tax records, they compare migrant children who move with their families to those stayed in higher-poverty areas. They find that children whose families had migrated to lower-poverty neighborhoods before they turned 13 experienced a 2.5 percentage point increase in college attendance probability at age 18-20 compared to those living in less advantaged areas. In addition, migrant children before the age of 13 attend better colleges in their destination areas. However, moving between 13 and 18 years old has a negative impact on college attendance at ages 18-20. Their interpretation is that moving at older ages to a new environment might disrupt social networks and adversely affect child development. Chetty and Hendren (2018) show that US children whose families had moved to commuting zones with higher college attendance rates are more likely to attend college at ages 18-23. Their identifying assumption is that the selection effects of families into better or worse areas are orthogonal with the child's age at the move. The authors provide evidence supporting the plausibility of this assumption by controlling for household fixed effects, time-varying observable characteristics within the family, and exogenous displacement shocks. They find that the incomes and college attendance rate in adulthood of migrant children converge to the ones of children of permanent residents in the commuting zones of destination at a rate of 4 percent per year. Using census data of 27 African countries and following the identification strategy of Chetty and Hendren (2018), Alesina et al. (2021) rely on the variation in the timing of children's migration across regions to assess the regional childhood

exposure effects. They find that children whose families migrate before they aged 12 to areas with higher educational intergenerational mobility are more likely to achieve primary school. Educational mobility is defined as the probability of a kid born to parents without any education to complete primary school. According to the authors, the region of destination matters through the higher quality of infrastructures and schools as well as the spatial sorting of families.

The purpose of the paper is to investigate the effect of children's migration on educational outcomes in South Africa. First, using the five waves of the National Income Dynamic Study (NIDS) individual panel data from 2008 to 2017, we compare the school progression and educational attainment of migrant children to non-migrant children by exploiting the within variation of migration among siblings. Second, the heterogeneity of the effect of children's migration by gender and the type of move (rural-rural, urban-urban, rural-urban, urban-rural) is examined. The rationale behind looking at the effect of the different types of migration is that a rural-urban move and a within-urban move might be associated with better educational opportunities through access to better schools while a within-rural and urban-rural migration does not seem to provide such opportunities. Basically, we would compare the school quality between the locality of origin and the destination area of migrant children. Unfortunately, we do not have access to the NIDS secure data, which contain information on children's schools (the name and geolocation of school, the number of learners and teachers, and the school quintile). Third, we focus on the migration status and the timing of the move of mothers by comparing children who migrate with their moms to non-migrant children. We hypothesize that moving with your biological mother might be positively associated with the child's cognitive development and school performance. Moreover, migrant mothers might gain increased income since moving to urban areas seem to provide greater job opportunities allowing mothers to fund the schooling fees of their children. Non-migrant children and those moving away from their mothers might deal with emotional stress due to parental absence, which could adversely affect educational outcomes. Nevertheless, the left-behind children might receive financial support from their mothers for education expenditure, for instance, leading to a likely lower adverse effect of being away from their biological mothers.

The contribution of this article to the literature is twofold. While most of existing studies focused on the effect of parental labour migration on children left behind, this paper goes further by analyzing the trajectory in terms of school progression and educational attainment of children who move with their mothers relative to non-migrant children. We also examine the heterogeneity of the effect of children's migration by the type of move in a context characterized by high spatial inequalities in the quality of schools. Furthermore, taking advantage of the indi-

vidual panel data over a 10-year period, we try to overcome the identification concerns such as unobserved individual heterogeneity and the endogeneity issue of child's move.

Examining the effect of children's migration on schooling outcomes raises the question of the selection issue on the one hand, and the challenge related to the definition of the counterfactual on another hand. Migration is potentially selective given that migrant children may have unobserved characteristics such as higher schooling and cognitive abilities compared to non-migrant children. To rule out this threat for identification, two empirical strategies are undertaken. The first one relies on a child panel sample and controls for individual unobserved heterogeneity and baseline household characteristics, whereas the second one based on a cross-section sample of siblings accounts for the mother's unobserved heterogeneity. Furthermore, non-migrant children might be left behind, and comparing these children to migrants might overestimate the effect of migration on educational outcomes due to the likely vulnerability associated with being left behind. To deal with this identification concern, I construct a control group composed of older children aged more than 18 years old at the time of maternal migration. Results suggest that children who moved at early ages with their mothers have a better school progression and are more likely to achieve grade 12.

The remainder of the paper is organized as follows. Section 2 provides background on school inequalities in South Africa. Section 3 presents the data and descriptive statistics and stresses the determinants of children's migration. Section 4 explains the identification strategies. Section 5 shows the main results and the heterogeneity of the effect by gender, type of move, and migration status of the mother. Section 6 discusses the robustness check analysis and concludes.

3.2 Data and descriptive statistics

This section explains the construction of the two samples of interest and presents descriptive statistics on migration and spatial inequalities in schools.

3.2.1 Data

This study uses individual panel data from the National Income Dynamics Study (NIDS) survey conducted in South Africa over a 10 year period (from 2008 to 2017). The baseline sample of the NIDS data is nationally representative and consisted of over 28,000 individuals in 7,300 households (Brophy et al. 2018). After the baseline survey in 2008, individuals are tracked in each wave and re-interviewed in 2010, 2012, 2014, and 2017. The five waves contain information on individual and household characteristics, and migration in four population groups: African (also

called black), Coloured, Asian/Indian, and White. The attrition rate is relatively low among the African population, about 14% over the period. The rationale behind focusing on black children is that these children are highly mobile relative to the other groups of the population, and their mobility is associated to a large extent with the mother's labour migration.

3.2.2 Sample

Two types of sample are constructed to examine the effect of children's migration on educational outcomes. The first sample is an individual unbalanced panel composed of black children aged 0-14 in the baseline survey in 2008 and living in a baseline household with two children and more. The children are observed at least two times over the 10 year period allowing us to examine changes in the school progression. Table 3.1 indicates the number of African children observed in each wave between 2008 and 2017. Keeping baseline households with two children and more allows us to exploit the within variation in the household by comparing the school progression of migrant children and non-migrant children living in the same initial household. Among the 5803 black children in the sample (aged 0-14 years old in 2008, and 8-24 years in 2017), 91% of children are observed in all five waves.

The second sample is an individual cross-section data and includes 1915 siblings aged 18 years old and more in 2017 as shown in Figure 3.1. The sample is restricted to siblings aged 18 and more in 2017 due to the outcome of interest, which is the grade 12 achievement. In South Africa, the theoretical age to complete grade 12 is 18. The mothers of siblings have at least two children as the empirical strategy exploits the within variation among siblings. The cross-section sample is constructed to assess the effect of migration on grade 12 achievement, which is a time-invariant outcome.

3.2.3 Measurement of educational outcomes

Two types of educational outcomes are constructed for the unbalanced individual panel and the cross-section data respectively. First, the school gap is the difference between the current grade of the child and the grade that the child should attend if he/she did not experience a grade repetition, drop-out, or late entry in school. It is computed for children aged 7 years old and more in each wave from 2008 to 2017. The educational attainment is coded zero for children aged 7 years old and more who are not currently enrolled in school. A negative value of the school gap means a school delay or a drop-out, a value of zero means that the child currently attends the right grade, and a positive value means an educational advance. In South Africa,

schooling is compulsory from 7 to 15 years old corresponding to grade 1 to grade 9. However, the education system is characterized by high-grade repetition and drop-out rates, especially in secondary school (grades 8 to 12) as stressed by Branson, Hofmeyr, and Lam (2014) and Branson and Lam (2017). Using the first three waves of the NIDS data, Branson and Lam (2017) show that only 40% of girls and 30% of boys who were enrolled in grade 9 in 2008 had completed grade 12 by 2012 without grade repetition. Figure 3.2 shows that the enrolment rate of black individuals is above 90% from 7 to 15 years old and decreases dramatically after the schooling compulsory age. Figure 3.3 presents the evolution of the school gap for black children between 2008 and 2017. The proportion of children experiencing a school delay is relatively important, about 65% on average over the period.

The second educational outcome is a time-invariant variable. It is a dummy variable equal to one if the child achieved grade 12, and zero otherwise. Given that half of the children completed grade 12 with delay as shown in Figure 3.4, we also construct two dummies indicating children who achieved grade 12 without delay (at the age 18), and those who completed the grade with delay (at age 19 and more).

3.2.4 Migration measurement and spatial inequalities in school quality

Migration is defined as a change in the place of residence between waves. In the baseline survey, children in the panel sample are non-migrants. The migration event occurs after 2008. Table 3.2 presents descriptive statistics on the migration status of black children and their mothers during the 10 year period. Nearly 70% of children who have a migrant mother moved residence between 2008 and 2017 suggesting that the child's move is strongly associated with maternal migration. Looking at the type of move of children over the period, Table 3.3 shows that moves within rural and within urban areas are relatively important compared to rural-urban or urban-rural migrations. These figures are in line with the migration literature in South Africa which stressed the importance of within rural and urban area migrations (Garlick, Leibbrandt, and Levinsohn 2016; Hall and Posel 2019). Figure 3.5 presents the household residency status of black children's parents between 2008 and 2017. During the period 2008-2017, 67% of children are co-residents with their mothers on average while only 24% are co-residents with their fathers on average. Given that children are more likely to co-reside with their mothers than their fathers, the migration of the mother might lead to the child's move. In South Africa, the proportion of black children who co-reside with their both parents is very low: 23% and 17% of children co-reside with their mothers and fathers in 2008 and 2017 respectively. Table 3.4 reports summary statistics on children in the panel sample. During the period 2008-2017, 42% of children

moved the place of residence and they migrated at age 10 on average. The number of children aged under 15 in the baseline household averaged 4.

To analyze spatial inequalities in schools, I downloaded school administrative data from the website of the Department of Basic Education (DBE) of South Africa from 2013 to 2017. The administrative data contain numerous information such as the number of learners and teachers per school, the school quintile, the geolocation (latitude and longitude) of schools, the school status (open or closed), the school sector (public or private), the school level (primary, secondary, combined). Schools are classified by quintile in South Africa based on the school's neighbourhood income, employment rate, and literacy levels calculated within the province from the last census (Branson and Lam 2017). The budget allocation depends on the school quintile ranking, the poorest schools receive more resources per pupil. In 2017, there were 22 463 public schools (primary, secondary, and combined) in the country. Figure 3.6 shows the distribution of schools of quintile Q5 (i.e., the 20% richest schools) in 2017 at the district level across the country. The five districts with the highest proportion of schools ranked in quintile Q5 are the City of Cape Town (44.6%), West Rand (32.5%), Ekurhuleni (32.2%), eThekweni (31.2%), and the City of Johannesburg (31.1%). Four of them are metropolitan districts indicating that better schools seem to be located in metropolitan areas. Figure 3.7 presents the distribution of schools ranked in quintile Q1 in 2017 at the district level across the country. The poorest schools are concentrated in traditional areas located in the provinces of Northern Cape, Limpopo, Free State, and KwaZulu-Natal.

3.2.5 Baseline characteristics and determinants of children's migration

Given that the decision of children's migration is not random, it is critical to examine the rationale behind the child's move and the characteristics of each group of children before discussing the identification strategy. Table 3.5 displays the baseline differences in observable characteristics in 2008 between movers and non-movers. The comparison of migrant and non-migrant children independently of the migration status of the mother indicates that movers are on average younger, more likely to live in rural areas and households where a member experienced serious illness in 2008 relative to non-movers. Migrant children are also more likely to have absent and educated parents in 2008. Looking at the baseline differences among children who have a migrant mother, descriptive statistics show some opposite patterns to the previous ones. Indeed, migrant children with a migrant mom and those moving with their mom are more likely to live in urban areas and wealthier households in 2008 and have less educated mothers compared to non-movers with a migrant mom. To analyze the determinants of children's migration,

we estimate a linear probability model. The outcome of interest is a dummy variable indicating whether the child has moved over the 10-year period, and the explanatory variables are the baseline characteristics measured in 2008 before the child's migration. Column 1 of Table 3.6 indicates that the child's move increases with the mother's education, living in households where a member experienced a serious illness, and decreases with the age of children. In addition, the probability to be a migrant child raises with the absence of parents in 2008 as shown in columns 2 and 3 of Table 3.6 suggesting that children may move to join their mothers.

3.3 Identification strategy

I use two econometric specifications to assess the effect of children's migration on educational outcomes. The first method exploits the variation within the baseline household by comparing the school progression of migrant and non-migrant children residing in the same initial household. The second approach relies on the variation within the mother by comparing the educational attainment between migrant and non-migrant siblings. The intuition behind estimating two different models is to analyze to what extent the results are consistent when the composition of children in the control group changes.

3.3.1 Fixed-Effects estimation – within household comparison

The effect of children's migration on school progression is estimated by using the following specification:

$$\Delta Schoolgap_{ih(t+1,t)} = \alpha_h + \beta M_{i(t-1)} + \lambda_t + \epsilon_{iht} \quad (3.1)$$

$\Delta Schoolgap_{ih(t+1,t)}$ is the variation of the school gap of children i in the baseline household h between waves ($Schoolgap_{ih(t+1)} - Schoolgap_{ih(t)}$); it can be interpreted as school progression.

$M_{i(t-1)}$: a dummy variable equal to 1 if the child i migrated at time $t-1$; at the baseline period in 2008, the dummy equal to zero for all children, the migration event occurred after 2008. The migration variable is lagged to ensure that the migration event is prior to the computation of the school progression.

α_h : the baseline household fixed effects which allow us to control for observed and unobserved

time-invariant heterogeneity in the initial household in 2008.

λ_t : the wave year dummies control for time-specific shocks that are common to all children.

The coefficient of interest is β . It measures the difference in school progression between migrant and non-migrant children who lived in the same baseline household in 2008.

The specification controls for individual and households of origin fixed effects. The individual fixed-effects¹ capture all time-invariant characteristics of the child such as ethnicity, gender, year of birth, cognitive and schooling abilities, and frailty allowing us to rule out a number of sources of endogeneity. For instance, if children with higher cognitive and schooling abilities are more likely to move, this could lead to a selection and bias the point estimates. Including the child fixed heterogeneity addresses to a large extent the selection issue. Furthermore, the baseline household fixed effects account for the initial household-level heterogeneity such as schooling preference, awareness of the importance of education, and public service delivery in the baseline location, which may affect both the educational outcome and the child's move. Thus, the empirical strategy exploits the variation within the initial household by comparing migrant and non-migrant children from the same household of origin. I compare siblings and relatives who were residing in the same initial household, and then I restrict the analysis to siblings. Comparing siblings and relatives in the baseline household raises the question of the comparability of children as the transmission of education and human capital investment may differ between siblings and other relatives. This leads us to restrict the comparison among siblings who were living together in 2008. To deal with the potential confounding factors such as time-varying shocks, which may affect both school progression and migration, I control for parental death between 2008 and 2017.

Although the identification strategy accounts for time-invariant child heterogeneity, fixed baseline household heterogeneity, and time-varying shocks, one may argue that the endogeneity of the child's move remains a threat to the identification due to child unobservable characteristics. An instrumental variable approach would be the ideal strategy to overcome the endogeneity concern. Basically, I would use an interaction variable between the birth order and the distance from the locality of origin to the metropolitan areas to instrument the children's migration. Metropolitan areas may offer better employment opportunities and schools. For example, the first-born child who was living in a baseline household relatively close to metropolitan localities

¹By construction, the children's individual fixed effects are accounted for in the specification due to the way the outcome of interest (school progression) is computed.

may have a higher probability to move. Unfortunately, I do not have access to the NIDS secure data, which contain information on the geolocation of children's households. Nevertheless, the set of fixed effect variables capture to a large extent the potential sources of endogeneity.

3.3.2 Cross-section estimation with mother fixed-effects

One limitation of the first empirical strategy is the control group as the non-migrant could be left behind children. Comparing children who move to join their mothers or co-migrate with their mothers to those left behind might overestimate the effect of migration on the schooling outcome due to the likely vulnerability of children in the control group. As explained above, left behind children may face emotional stress and a lack of care and adult supervision due to the absence of the biological mother, which could lead to worse educational outcomes. To overcome this identification issue, I construct a control group composed of older children aged more than 18 years old at the time of maternal migration as shown in Figure 3.1. Given that the theoretical age to complete Grade 12 is 18 in South Africa, these older children are not affected by the mother's move. Table A1 shows descriptive statistics on treated and control groups.

I assess the effect of children's migration on Grade 12 achievement by estimating the following econometric model:

$$AchievedG12_{im} = \alpha_m + \beta_1 Treat1_i + \beta_2 Treat2_i + X_i + \epsilon_{im} \quad (3.2)$$

AchievedG12_{im}: a dummy equal to 1 if the child *i* of mother *m* achieved Grade 12, and zero otherwise; it also indicates whether children achieved Grade 12 without delay (at the age 18), or completed it with delay (at age 19 and more) respectively.

Treat1_i: a dummy indicating whether children migrated between 0 and 17 years old.

Treat2_i: a dummy indicating whether the child is left behind, i.e. aged under 18 at the time of maternal migration and never move or moved after age 18.

α_m : is mother fixed effects which account for observed and unobserved time-invariant heterogeneity of the mother.

X_i : refers to the child's individual time-invariant characteristics such as gender, year of birth,

and birth order.

The coefficients of interest are β_1 and β_2 . They capture the effect of children's migration and being left behind children on Grade 12 achievement compared to the older ones aged more than 18 at the time of maternal migration respectively. The mother's individual heterogeneity controls for a number of time-fixed characteristics such as the level of education, the ability, the awareness and valuation of education. Taking into account the mother fixed effects allows us to exploit the variation within the mother by comparing the probability of Grade 12 achievement among siblings.

Even if the identification strategy controls for the child's observable characteristics and compares siblings from the same mother, there are still children's unobserved characteristics that might affect both children's migration and Grade 12 achievement. As underlined above, an instrumental variable method would be the ideal strategy to rule out the endogeneity issue but the data limitation constrains us to do so.

3.4 Results

3.4.1 Main results

Table 3.7 presents the estimates of the effect of children's migration on school progression. Panel A shows the results when the sample includes all children in the baseline household while Panel B displays the results when the sample is restricted to siblings who were living together in the same household in 2008. The first column of Table 3.7 refers to the overall effect of migration regardless of the timing of the move, and the second one reports the heterogeneity of the effect by the age at move. The rationale behind splitting the age at move into three categories is motivated by the structure of the education system in South Africa. Indeed, children should start Grade 1 at age 7 and complete primary school (Grade 7) at age 13. They should attend secondary school from the age 14 (Grade 8) to age 18 (Grade 18). The first category of ages 0-6 corresponds to the preschool age, the second one (7-13) to the primary school age, and the third one (14-18) to the secondary school age. Since, by construction, the school progression is computed for children aged 7 years old and more across waves, the variable indicating whether the child moved between the ages 0 and 6 is not lagged because this migration event occurred before the computation of the outcome of interest. All columns control for individual and baseline household fixed effects, and wave year dummies.

While column (1) in Panel A of Table 3.7 indicates a negative and not significant effect of

migration on school progression, column (2) in Panel A of Table 3.7 shows that children moved between the ages 0 and 13 have a better school progression compared to non-migrant children. The negative effect in column (1) of Table 3.7 is driven by children moved at ages 14-18 who experienced a lower educational progression compared to non-migrant. When the sample is restricted to siblings, the point estimates display a similar pattern. Moving at early ages is positively associated with school progression whereas migrating in later life is associated with worse educational outcomes relative to non-migrant as shown in Panel B of Table 3.7. According to the authors Chetty, Hendren, and Katz (2016), moving during adolescence to a new environment, in particular between the ages of 13 and 18 might disrupt social ties, which could adversely affect educational outcomes. Hence, the negative association between moving at ages 14-18 and school progression could reflect a likely disruption effect.

Table 3.8 reports the effect of children's migration and being left behind children on grade 12 achievement. The control group is older children who were not affected by maternal migration, i.e. those aged more than 18 years old at the time of the mother's migration. Panel A shows the overall effect of the two treatments, and Panel B presents the heterogeneity of the effect by the age at move and the left behind. The outcomes are ordered as follows: (1)-(4) probability to achieve grade 12 irrespective of the age of achievement, (2)-(5) probability to complete grade 12 without delay (at age 18), (3)-(6) probability to complete grade 12 with delay (at age 19 and more). All specifications account for the child's individual characteristics such as gender, year of birth, and birth order. In the last three columns of Table 3.8, I control for mother fixed effects. Results indicate that moving in early life, especially between the ages 7 and 13 is positively associated with the probability to complete grade 12 with delay as shown in column (3) of Panels A and B of Table 3.8. However, the positive association becomes statistically insignificant once the mother's individual heterogeneity is accounted for suggesting that the positive effect was driven by the mother's time-invariant unobserved characteristics. For instance, mothers may display preferences by investing more in children with higher schooling abilities. Overall, the estimates from the cross-section sample suggest that the selection effect among children matters.

3.4.2 Heterogeneity by gender and type of move

Table 3.9 presents the estimates of the effect of children's move on educational progression by gender. Results indicate that only boys who moved between the ages 0 and 13 experienced a positive school progression and a negative one if they moved at ages 14-18 as shown in column (2) of Table 3.9. I find no significant effect on educational progression for migrant girls. This difference in school progression between migrant boys and girls may come from the type of destination

area, the probability to move with your mother, or gender discrimination in human capital investment. For instance, if boys are more likely to move to urban areas with their moms, they could have better school performance. Looking at the probability to migrate to urban areas by gender, I find no significant difference between boys and girls. Furthermore, the baseline characteristics in Table 3.5 report that there is no gender difference in the probability of moving with the mother.

Figure 3.8 shows the effect of migration on school progression by the type of move. The point estimates report that children who moved at ages 0-6 within urban areas and those who moved between the ages 7 and 13 from rural to urban areas have a higher educational progression compared to non-migrant. One likely explanation is the better school quality in the locality of the destination since urban areas offer better schooling opportunities². However, moving at ages 14-18 is negatively associated with school progression for all types of migration suggesting a likely disruption effect as discussed above. Table 3.10 presents the heterogeneity of the effect of migration on educational progression by gender and type of move. Results reveal that the positive association between moving within urban areas as well as rural-urban move and school performance is driven by boys. Indeed, boys who migrated in early life to urban areas performed more at school compared to non-migrant as pointed out in columns (2) and (3) of Table 3.10. The magnitude of the coefficient is larger and more statistically significant for boys who moved from rural to urban areas at ages 7-13 suggesting a likely better access to high-quality schools. However, girls who moved from rural to urban areas at ages 7-13 experienced lower educational progression (column (3) of Table 3.10). Again, moving between the ages 14 and 18 is still negatively associated with school progression irrespective of the type of move.

Table 3.11 reports the estimates of the effect of migration on grade 12 achievement by gender from the cross-section sample³. Results indicate that sorting among siblings drives the point estimates since the coefficients become statistically insignificant and smaller in magnitude once the mother's unobserved heterogeneity is accounted for. Migration and being left behind children are positively associated with grade 12 achievement only for girls as shown in columns (1) and (3) of Table 3.11, but the effect disappears once we control for the mother's fixed effects in columns (4) and (6) of Table 3.11. To sum up, sorting seems considerable leading us to interpret our results with caution.

²See Appendix A6 for more details.

³The information on the type of the area of residence of origin (rural or urban) is unavailable for individuals who moved before 2008. Thus, the heterogeneity by the type of move is not examined in the cross-section sample due to the data limitation.

3.4.3 Children's move and maternal migration

Table 3.12 shows the heterogeneous effects of children's move on educational progression by the migration status of the mother. Children who moved away from their mother or moved independently of the migration status of the mother are potentially foster children while those who migrated with their mother are still living with her. Moving away from your biological mother or being a foster child may be associated with greater emotional stress, a lack of care and adult supervision, and more domestic chores in the host household. Results suggest that migrating with your mother is positively associated with school progression whereas moving away from her leads to worse educational performance (column (1) of Table 3.12), which seems consistent with our hypothesis. Looking at the heterogeneity of the effect by the child's age at move with her mother, I find that children who migrated between the ages 0 and 6 with their mom experience better school progression, and those who moved at ages 14-18 away from their mom progress less relative to non-migrant (column (2) of Table 3.12). Compared to the results from Zimmerman (2003) on the positive effect of black child fostering on school attendance, our analysis provides new evidence indicating that moving away from your biological mother is negatively associated with school progression.

Table 3.13 presents the effect of the child's move on grade 12 achievement taking into account the maternal migration. Overall, estimates from the cross-section sample are in line with the panel ones even if sorting appears important. Results show that moving with your mother, especially at ages 7-13, increases the probability to complete grade 12 by 23.4 percentage points (columns (1) and (4) of Table 3.13). The coefficients remain statistically significant and relatively stable in magnitude when the mother fixed effects are controlled for. To sum up, moving with your biological mother seems to be associated with better school progression and a higher probability to complete grade 12.

Results from the panel sample indicate that children's migration is positively associated with educational progression. However, Equation (1) does not control for any time-varying child characteristics raising the omitted variable concern. One may argue that there are time-variant shocks, which could affect both the child's move and school performance. For instance, a parental death during the 10-year period may lead to children's migration and adversely affect the progression at school. Failing to control for such confounding factors could bias the point estimates. To rule out this potential threat for the identification strategy, I control for a parental death between 2008 and 2017. Results from Tables A2, A3, A4, and A5 reveal that the coefficients remain statistically significant and stable in magnitude suggesting that our point estimates are

not driven by a death shock.

3.5 Conclusion

This paper investigates the effect of children's migration on educational outcomes in South Africa. Exploiting the variation within the baseline households and within the mothers respectively, I compare migrant and non-migrant siblings. Results from the panel sample suggest that children's migration is positively associated with school progression. Analyzing the heterogeneity of the effects by gender and type of move, I find that boys who moved to urban areas have a better educational progression compared to non-migrant. Looking at the child's move in relation to maternal migration, I show that children who moved with their mother performed more at school while moving away from your mother is negatively associated with school progression. However, the positive effect of migration on school progression might be driven by the likely vulnerability of children in the control group as it includes left-behind children. Indeed, when I change the control group by considering older children who were not affected by maternal migration in the cross-section sample, I do not find any more a positive and statistically significant effect of migration on schooling outcomes. This suggests that the point estimates from the panel sample should be interpreted with caution due to the likely vulnerability of the group control. Furthermore, results from the cross-section sample indicate that sorting among siblings matters since the coefficients become statistically insignificant and smaller in magnitude once the mother's unobserved heterogeneity is accounted for.

Another limitation of the study is that I do not have access to the NIDS secure data, which contain information on the geolocation of children's households and schools as well as the number of learners and teachers, and the school quintile. It would be useful to examine the school quality between the locality of origin and the destination of migrant children. Even if our results present some caveats, they suggest that moving with your biological mother may lead migrant children to have a better school progression. An interesting follow-up of this paper could be to explore the mechanisms at work by comparing the income, education expenditures, and child labor between the households of migrant children and those who are left behind over time. In addition, it would be interesting to shed light on changes in the mother's employment status and level of income before and after the migration event.

REFERENCES

- Alesina, Alberto et al. (2021). “Intergenerational mobility in Africa”. *Econometrica* 89.1, pp. 1–35.
- Amuedo-Dorantes, Catalina and Susan Pozo (2010). “Accounting for remittance and migration effects on children’s schooling”. *World development* 38.12, pp. 1747–1759.
- Antman, Francisca M (2011). “The intergenerational effects of paternal migration on schooling and work: What can we learn from children’s time allocations?” *Journal of Development Economics* 96.2, pp. 200–208.
- (2012). “Gender, educational attainment, and the impact of parental migration on children left behind”. *Journal of Population Economics* 25.4, pp. 1187–1214.
- Branson, Nicola, Clare Hofmeyr, and David Lam (2014). “Progress through school and the determinants of school dropout in South Africa”. *Development Southern Africa* 31.1, pp. 106–126.
- Branson, Nicola and David Lam (2017). “The impact of the no-fee school policy on enrolment and school performance: Evidence from NIDS Waves 1-3”.
- Brophy, Timothy et al. (2018). “National income dynamics study panel user manual”. *Technical Note Release*.
- Case, Anne and Angus Deaton (1999). “School Inputs and Educational Outcomes in South Africa”. *The Quarterly Journal of Economics* 114.3, pp. 1047–1084.
- Chetty, Raj and Nathaniel Hendren (2018). “The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects”. *The Quarterly Journal of Economics* 133.3, pp. 1107–1162.

- Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz (2016). “The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment”. *American Economic Review* 106.4, pp. 855–902.
- Chyn, Eric and Lawrence F Katz (2021). “Neighborhoods matter: Assessing the evidence for place effects”. *Journal of Economic Perspectives* 35.4, pp. 197–222.
- Garlick, Julia, Murray Leibbrandt, and James Levinsohn (2016). *Individual migration and household incomes*. Tech. rep. National Bureau of Economic Research.
- Hall, Katharine and Dorrit Posel (2019). “Fragmenting the family? The complexity of household migration strategies in post-apartheid South Africa”. *IZA Journal of Development and Migration* 10.2, pp. 1–20.
- (2020). *What does labour migration mean for families? Children’s mobility in the context of maternal migration*. Migrant Labour after Apartheid: Continuities and Change. Cape Town: HSRC Press.
- Köhler, Timothy (2020). “Class size and learner outcomes in South African schools: The role of school socioeconomic status”. *Development Southern Africa*, pp. 1–25.
- Madhavan, Sangeetha et al. (2012). “Child mobility, maternal status, and household composition in rural South Africa”. *Demography* 49.2, pp. 699–718.
- Meng, Xin and Chikako Yamauchi (2017). “Children of migrants: The cumulative impact of parental migration on children’s education and health outcomes in China”. *Demography* 54.5, pp. 1677–1714.
- Spaull, Nicholas (2013). “Poverty & privilege: Primary school inequality in South Africa”. *International Journal of Educational Development* 33.5, pp. 436–447.
- Zhang, Hongliang et al. (2014). “Does parental absence reduce cognitive achievements? Evidence from rural China”. *Journal of Development Economics* 111, pp. 181–195.
- Zimmerman, Frederick J (2003). “Cinderella goes to school the effects of child fostering on school enrollment in South Africa”. *Journal of Human Resources* 38.3, pp. 557–590.

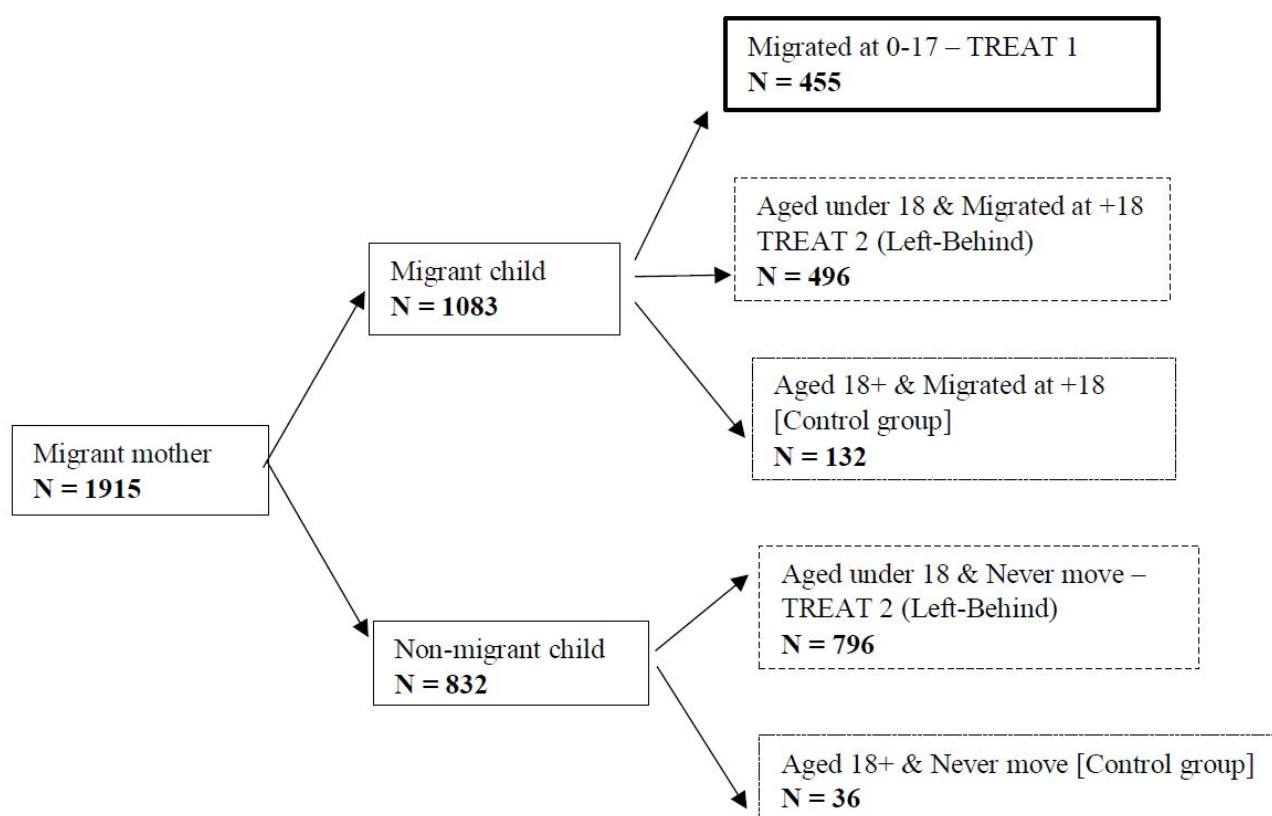


Figure 3.1: Cross-section sample

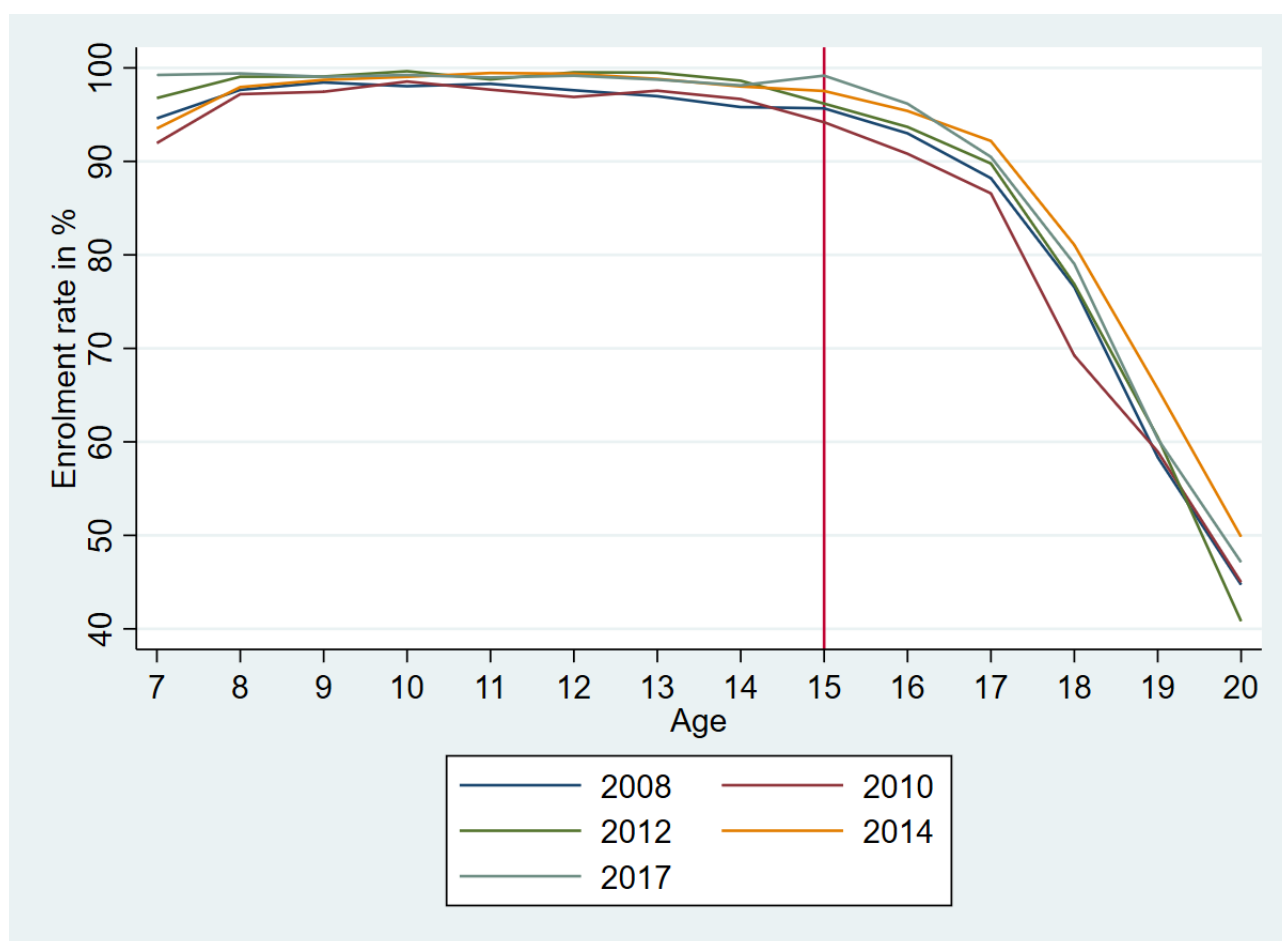


Figure 3.2: Enrolment rate of black individuals aged 7-20 in each wave from 2008 to 2017

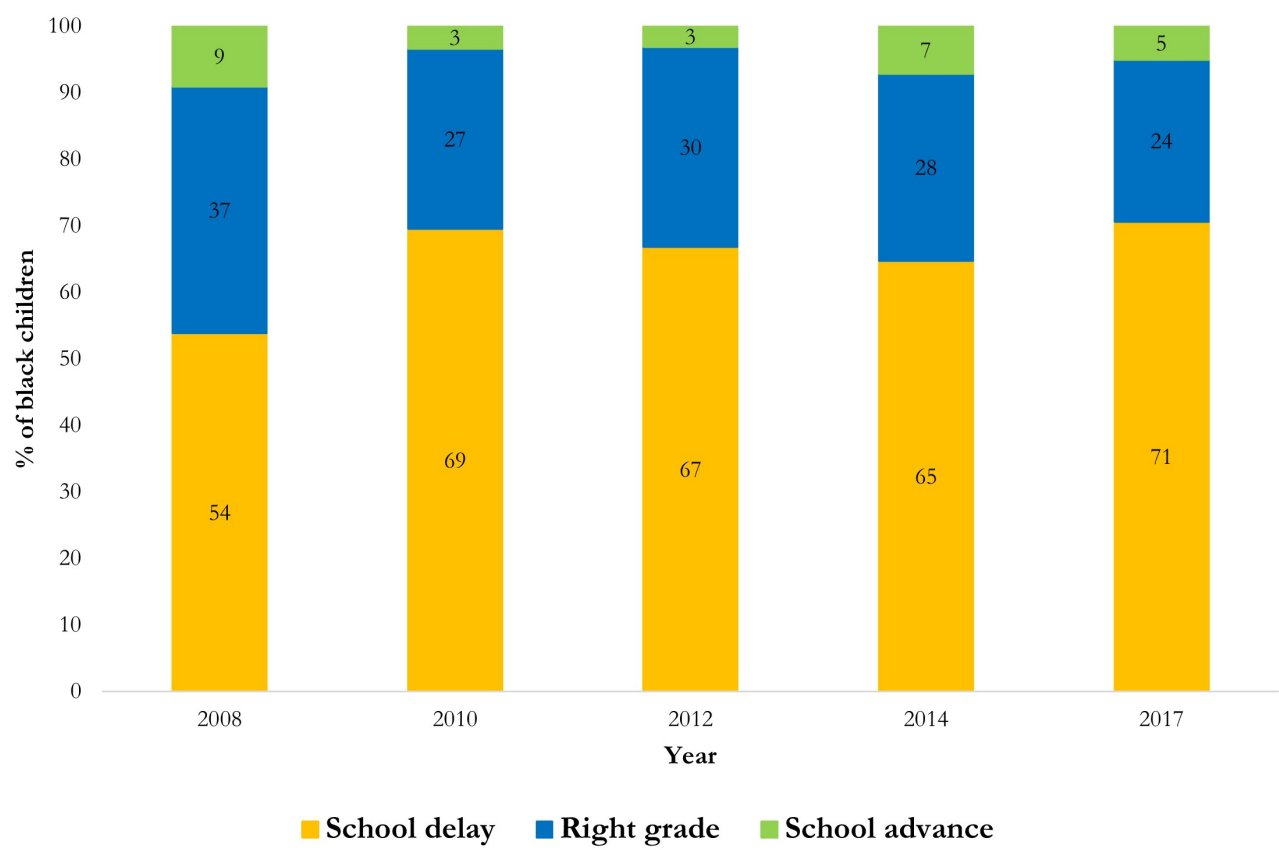


Figure 3.3: School gap of black children 2008-2017

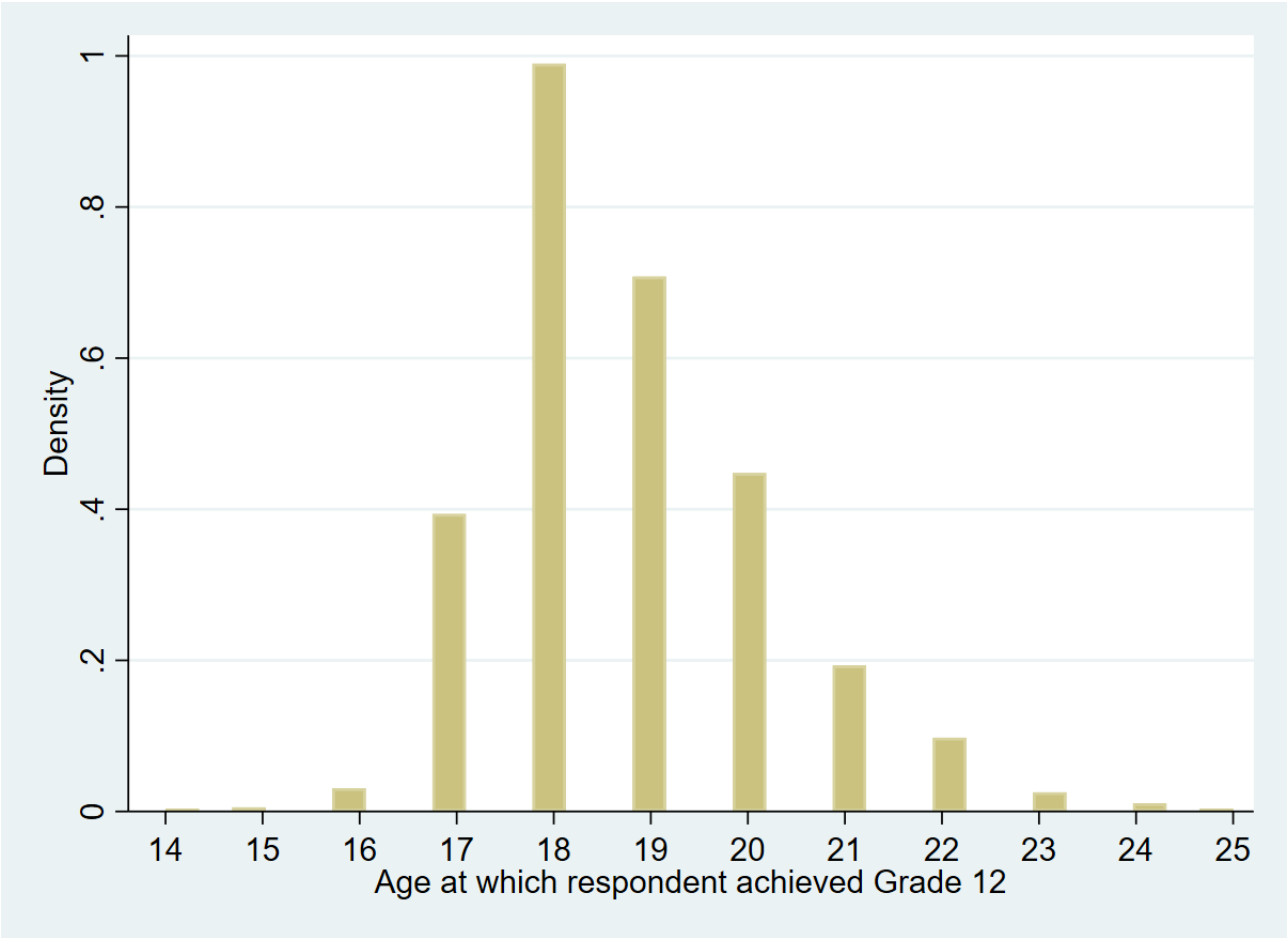


Figure 3.4: Age at Grade 12 achievement

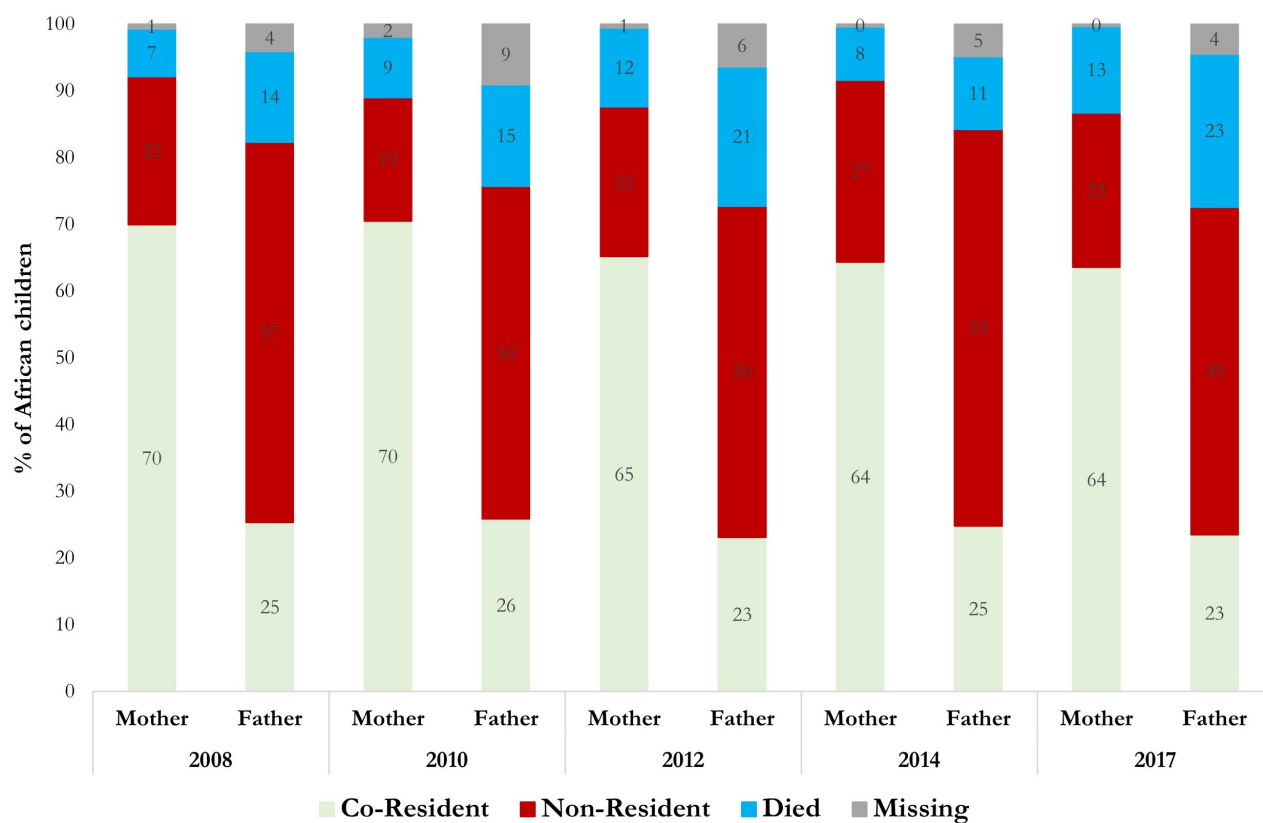


Figure 3.5: Household residency status of black children's parents 2008-2017

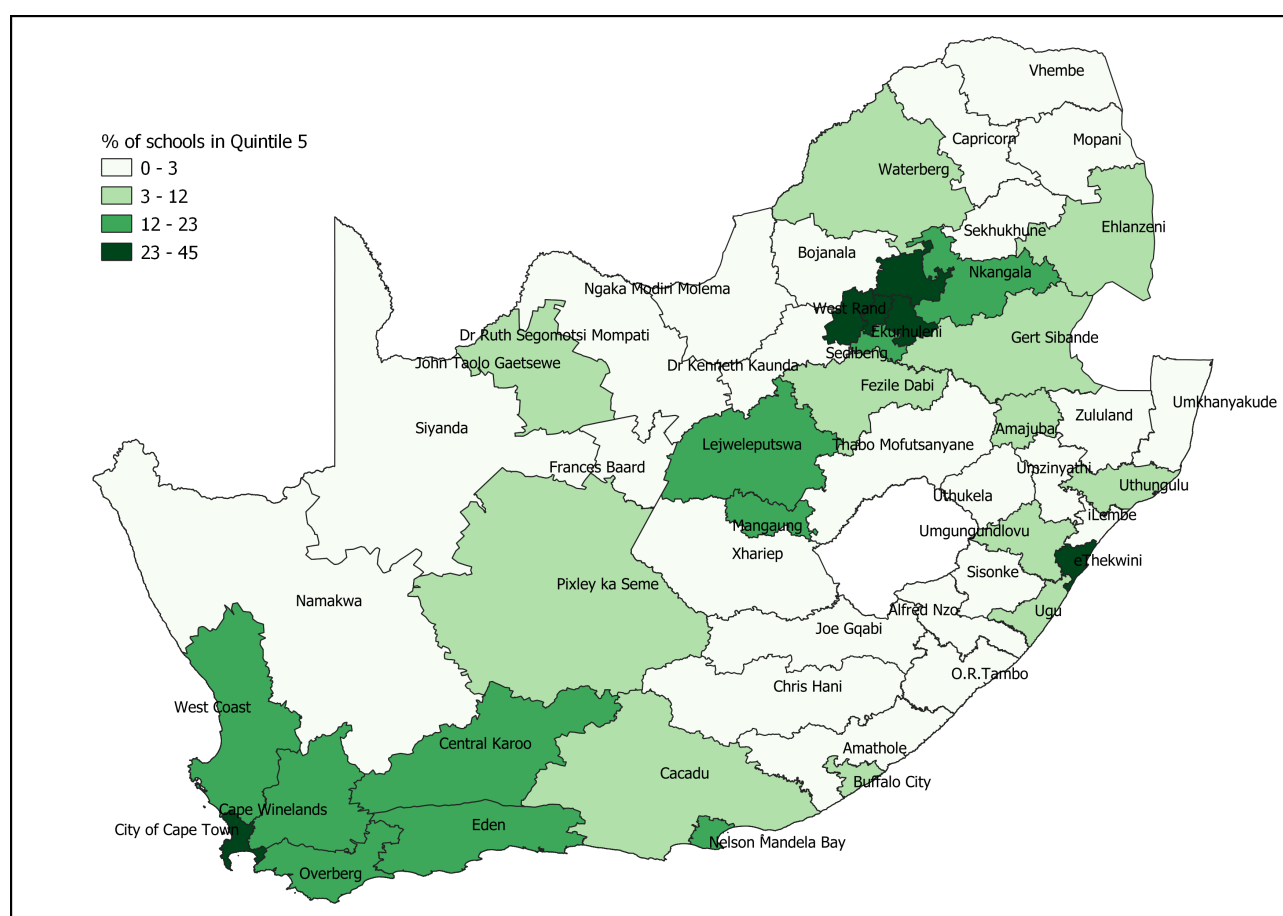


Figure 3.6: Proportion of schools in quintile Q5 at the district level in 2017

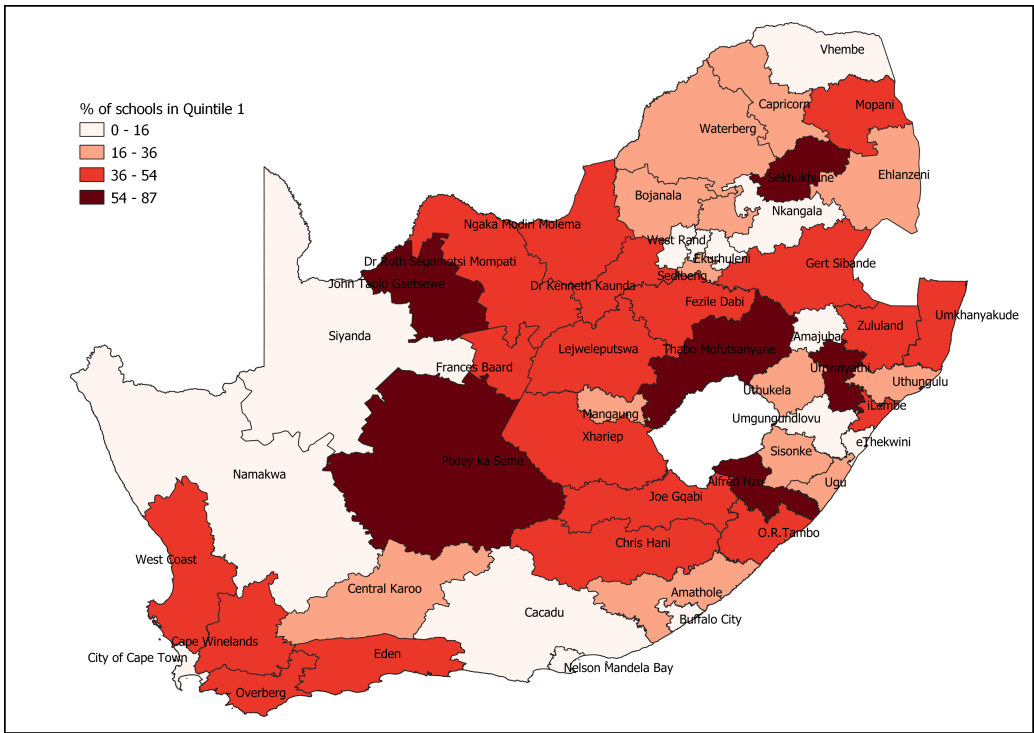


Figure 3.7: Proportion of schools in quintile Q1 at the district level in 2017

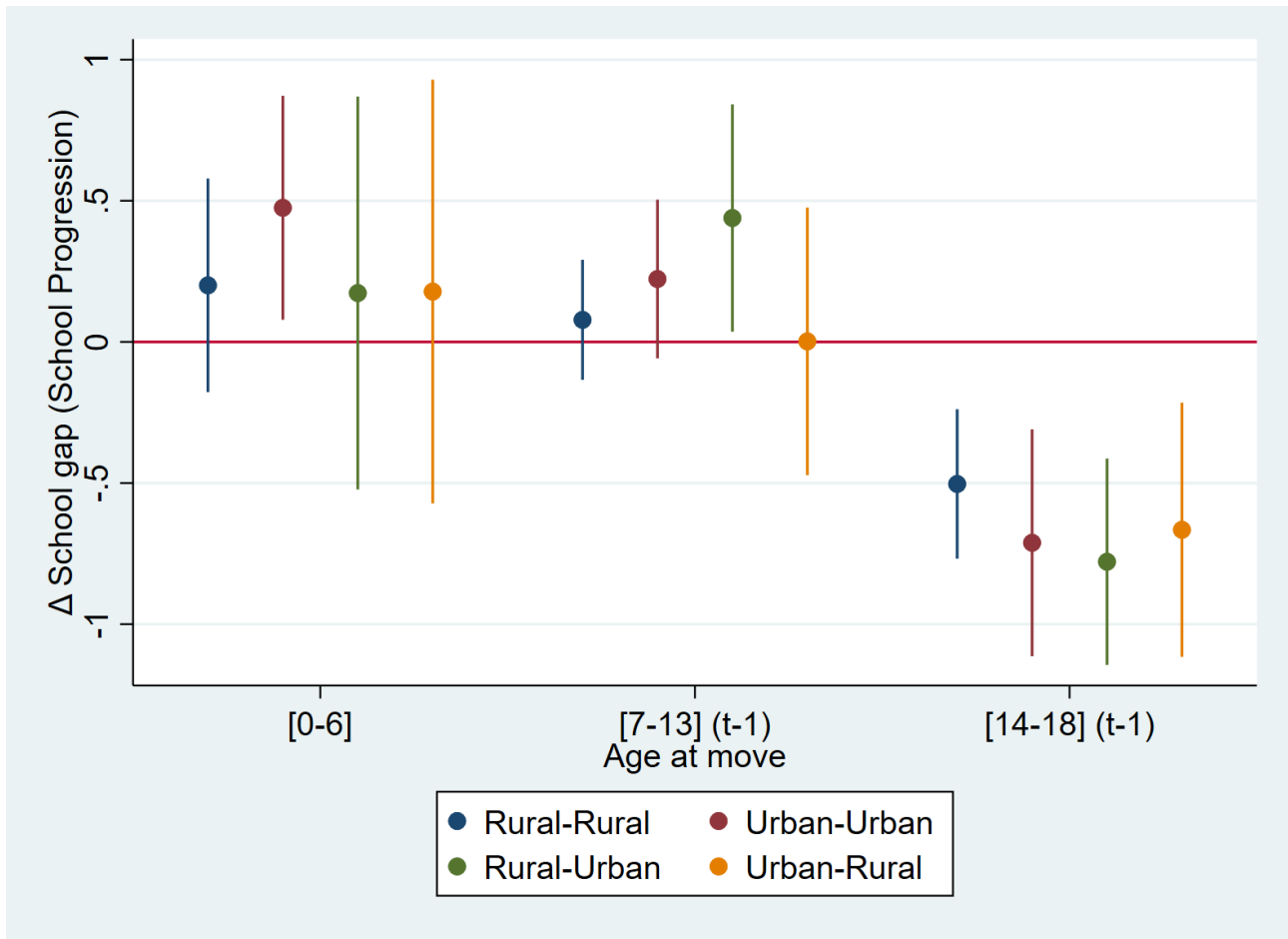


Figure 3.8: Type of move and School progression

Table 3.1: Number of black children observed in each wave (2008-2010-2012-2014-2017)

	Frequency	Percentage
Observed in 5 waves	5280	90.99
Observed in 4 waves	397	6.84
Observed in 3 waves	91	1.57
Observed in 2 waves	35	0.60
Total	5803	100

Table 3.2: Migration status of black children and their mothers 2008-2017

	Non-migrant child	Migrant child	Total
Non-migrant mother	2657	903	3560
	74.63	25.37	100
Migrant mother	690	1553	2243
	30.76	69.24	100
Total	3347	2456	5803

Table 3.3: Type of migration of black children 2008-2017

Type of migration	Frequency	Percentage
Rural-Rural	1089	44.34
Urban-Urban	622	25.33
Rural-Urban	466	18.97
Urban-Rural	279	11.36
Total	2456	100

Table 3.4: Descriptive Statistics

	Mean	Std. Deviation	Min	Max	N
Girl	0.49	(0.50)	0	1	5803
Year of birth	2000	(4.19)	1993	2008	5803
Age in 2008	6.60	(4.18)	0	14	5803
Age in 2017	15.55	(4.11)	8	24	5793
Migrant	0.42	(0.49)	0	1	5803
Year of migration	2011	(2.24)	2009	2017	2449
Age at migration	10.50	(4.55)	1	18	2449
Rural-Rural migration	0.19	(0.39)	0	1	5803
Urban-Urban migration	0.11	(0.31)	0	1	5803
Rural-Urban migration	0.08	(0.27)	0	1	5803
Urban-Rural migration	0.05	(0.21)	0	1	5803
School gap in 2017	-1.82	(2.10)	-12	2	5665
Number of children (0-14) in the baseline HH 2008	3.66	(1.70)	2	12	5803

Table 3.5: Baseline characteristics in 2008 between migrant and non-migrant children

	Non-migrant (a)	Migrant (b)	Migrant & migrant mum(c)	Non-migrant & migrant mum(d)	Migrant & moved with mum(e)	Diff (a)-(b)	Diff (d)-(c)	Diff (d)-(e)
Girl	0.48	0.5	0.48	0.46	0.49	-0.02	-0.02	-0.03
Age in 2008	6.77	6.36	5.48	5.99	4.89	0.42***	0.51***	1.10***
Area of residence 2008: Rural	0.69	0.71	0.68	0.76	0.6	-0.02*	0.08***	0.16***
Number of HH residents in 2008	7.21	7.23	7.32	7.44	7.35	-0.02	0.13	0.09
HH monthly food expenditure per capita in 2008	129.11	129.03	133.37	124.97	140.71	0.08	-8.4	-15.75**
HH monthly non-food expenditure per capita in 2008	199.92	193.73	206.34	140.38	266.04	6.19	-65.96**	-125.66***
Mother's education: at least primary	0.2	0.24	0.15	0.31	0.03	-0.05***	0.16***	0.28***
Father's education: at least primary	0.37	0.4	0.43	0.45	0.44	-0.03**	0.03	0.01
Biological mother absent 2008	0.21	0.31	0.22	0.46	0.05	-0.11***	0.24***	0.42***
Biological father absent 2008	0.66	0.75	0.72	0.84	0.66	-0.09***	0.12***	0.19***
Biological father died (YoD<=2008)	0.15	0.12	0.1	0.14	0.09	0.02***	0.04***	0.05***
HH member experienced serious illness/injury 2006-2008	0.03	0.05	0.04	0.03	0.05	-0.01**	0	-0.02*
Enrolled in school in 2008 (aged 7-14 in 2008)	0.98	0.98	0.97	0.99	0.98	0	0.02*	0.02*
Grade repetition in 2008 (aged 7-14 in 2008)	0.3	0.3	0.28	0.29	0.28	0	0.01	0.01
School gap in 2008 (aged 7-14 in 2008)	-0.75	-0.77	-0.73	-0.7	-0.58	0.02	0.03	-0.12

***p<0.01, **p<0.05, *p<0.1

Table 3.6: Determinants of children's migration

	(1)	(2)	(3)
Girl	0.017 (0.013)	0.017 (0.015)	0.020 (0.022)
Age in 2008	-0.006*** (0.002)	-0.007*** (0.002)	0.005 (0.005)
Area of residence 2008: Rural	0.020 (0.015)	0.004 (0.017)	0.016 (0.025)
Number of HH residents in 2008	-0.002 (0.002)	-0.002 (0.003)	-0.003 (0.004)
HH monthly food expenditure per capita in 2008 (log)	-0.015 (0.012)	-0.008 (0.014)	-0.029 (0.021)
HH monthly non-food expenditure per capita in 2008 (log)	-0.000 (0.006)	0.002 (0.006)	0.006 (0.010)
Mother's education: at least primary	0.071*** (0.016)	-0.036 (0.033)	-0.023 (0.045)
Father's education: at least primary	0.022 (0.013)	0.012 (0.018)	0.034 (0.027)
HH member experienced serious illness/injury 2006-2008	0.081** (0.034)	0.109*** (0.040)	0.099* (0.059)
Biological mother absent 2008		0.154*** (0.031)	0.189*** (0.040)
Biological father absent 2008		0.052** (0.020)	0.071** (0.029)
School gap in 2008 (aged 7-14 in 2008)			0.001 (0.010)
N	5803	4561	2037

Linear Probability Model (LPM). Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 3.7: Children's migration and School progression

	Δ School gap (School progression)	
	(1)	(2)
Panel A: all children in the baseline household in 2008		
Migrant (t-1)	-0.022 (0.029)	
Moved at [0-6]		0.334*** (0.051)
Moved at [7-13] (t-1)		0.223*** (0.035)
Moved at [14-18] (t-1)		-0.609*** (0.042)
N	14620	14620
Number of household FE	1884	1884
Panel B: siblings in the baseline household in 2008		
Migrant (t-1)	-0.137** (0.063)	
Moved at [0-6]		0.303** (0.122)
Moved at [7-13] (t-1)		0.142* (0.076)
Moved at [14-18] (t-1)		-0.614*** (0.085)
N	5120	5120
Number of household FE	752	752
Baseline household FE	Yes	Yes
Wave year FE	Yes	Yes

School gap is computed for children aged 7 years old and more in each wave from 2008 to 2017. Δ school gap is the difference in school gap between waves (2010-2008 ; 2012-2010 ; 2014-2012; 2017-2014). Δ school gap can be interpreted as school progression. Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 3.8: Migration and Grade 12 achievement

	Achieved G12	Achieved G12 at 18	Achieved G12 at 19+	Achieved G12	Achieved G12 at 18	Achieved G12 G12 at 19+
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Overall effect						
Treat 1	0.017 (0.046)	-0.025 (0.035)	0.061* (0.037)	0.019 (0.070)	0.008 (0.054)	0.056 (0.061)
Treat 2	0.043 (0.042)	0.003 (0.031)	0.036 (0.033)	0.053 (0.070)	-0.037 (0.053)	0.095 (0.061)
Girl	0.102*** (0.022)	0.109*** (0.017)	-0.004 (0.017)	0.093*** (0.025)	0.088*** (0.019)	0.016 (0.022)
N	1915	1915	1915	1915	1915	1915
Panel B: Heterogeneity by age						
Moved at [0-6]	0.012 (0.059)	-0.023 (0.045)	0.039 (0.047)	-0.062 (0.096)	0.011 (0.073)	-0.026 (0.084)
Moved at [7-13]	0.017 (0.053)	-0.038 (0.040)	0.088** (0.042)	-0.061 (0.084)	-0.039 (0.064)	0.040 (0.073)
Moved at [14-17]	0.029 (0.059)	-0.005 (0.044)	0.049 (0.047)	0.080 (0.083)	0.057 (0.063)	0.069 (0.072)
Left-Behind at [0-6]	0.047 (0.042)	0.004 (0.032)	0.040 (0.033)	0.011 (0.088)	-0.037 (0.067)	0.065 (0.077)
Left-Behind at [7-13]	0.060 (0.056)	0.024 (0.042)	0.028 (0.044)	0.031 (0.081)	-0.035 (0.061)	0.068 (0.070)
Left-Behind at [14-17]	-0.039 (0.072)	-0.050 (0.054)	-0.003 (0.057)	0.054 (0.086)	-0.049 (0.065)	0.102 (0.075)
Girl	0.104*** (0.022)	0.110*** (0.017)	-0.004 (0.018)	0.093*** (0.025)	0.090*** (0.019)	0.014 (0.022)
N	1915	1915	1915	1915	1915	1915
Year of birth FE & Birth order	Yes	Yes	Yes	Yes	Yes	Yes
Mother FE	No	No	No	Yes	Yes	Yes
Number of mother FE				722	722	722

Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 3.9: Children's migration and School progression - Heterogeneity by gender

	Δ School gap (School progression)	
	(1)	(2)
Siblings in the baseline household in 2008		
Migrant (t-1)	-0.080 (0.078)	
Migrant (t-1) \times Girl	-0.114 (0.094)	
Moved at [0-6]		0.343** (0.149)
Moved at [0-6] \times Girl		-0.065 (0.216)
Moved at [7-13] (t-1)		0.182* (0.099)
Moved at [7-13] (t-1) \times Girl		-0.074 (0.125)
Moved at [14-18] (t-1)		-0.567*** (0.119)
Moved at [14-18] (t-1) \times Girl		-0.095 (0.161)
N	5120	5120
Number of household FE	752	752
Baseline household FE	Yes	Yes
Wave year FE	Yes	Yes

Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 3.10: Children's migration and School progression - Heterogeneity by gender and type of move

	Δ School gap (School progression)			
	Rural-Rural move	Urban-Urban move	Rural-Urban move	Urban-Rural move
	(1)	(2)	(3)	(4)
Siblings in the baseline household in 2008				
Moved at [0-6]	0.195 (0.239)	0.555** (0.263)	0.354 (0.362)	0.247 (0.462)
Moved at [0-6] \times Girl	0.031 (0.347)	-0.138 (0.335)	-0.185 (1.140)	-0.079 (0.775)
Moved at [7-13] (t-1)	0.117 (0.156)	0.234 (0.164)	1.052*** (0.344)	-0.128 (0.289)
Moved at [7-13] (t-1) \times Girl	-0.091 (0.193)	0.024 (0.217)	-0.846** (0.370)	0.538 (0.477)
Moved at [14-18] (t-1)	-0.421** (0.175)	-0.839** (0.326)	-0.597** (0.268)	-1.106*** (0.334)
Moved at [14-18] (t-1) \times Girl	-0.185 (0.263)	0.190 (0.408)	-0.361 (0.356)	0.784* (0.444)
N	4067	3840	3656	3531
Number of household FE	578	576	521	516
Baseline household FE	Yes	Yes	Yes	Yes
Wave year FE	Yes	Yes	Yes	Yes

Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 3.11: Migration and Grade 12 achievement - Heterogeneity by gender

	Achieved G12	Achieved G12 at 18	Achieved G12 at 19+	Achieved G12	Achieved G12 at 18	Achieved G12 G12 at 19+
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Overall effect						
Treat 1	-0.093 (0.066)	-0.064 (0.050)	0.017 (0.052)	-0.001 (0.093)	0.005 (0.071)	0.050 (0.081)
Treat 1 \times Girl	0.206** (0.088)	0.072 (0.066)	0.084 (0.069)	0.033 (0.103)	0.005 (0.078)	0.010 (0.089)
Treat 2	-0.064 (0.059)	-0.005 (0.044)	-0.035 (0.047)	0.024 (0.089)	-0.024 (0.068)	0.058 (0.077)
Treat 2 \times Girl	0.203** (0.080)	0.013 (0.060)	0.136** (0.063)	0.050 (0.093)	-0.024 (0.070)	0.066 (0.080)
N	1915	1915	1915	1915	1915	1915
Panel B: Heterogeneity by age						
Moved at [0-6]	-0.123 (0.082)	-0.068 (0.062)	-0.028 (0.065)	-0.100 (0.120)	-0.024 (0.091)	-0.029 (0.104)
Moved at [0-6] \times Girl	0.259** (0.111)	0.090 (0.084)	0.124 (0.088)	0.070 (0.130)	0.068 (0.099)	0.000 (0.113)
Moved at [7-13]	-0.065 (0.077)	-0.065 (0.058)	0.073 (0.061)	-0.038 (0.108)	-0.020 (0.082)	0.066 (0.094)
Moved at [7-13] \times Girl	0.155 (0.102)	0.049 (0.077)	0.031 (0.081)	-0.040 (0.118)	-0.040 (0.090)	-0.037 (0.102)
Moved at [14-17]	-0.090 (0.086)	-0.053 (0.065)	-0.011 (0.068)	0.029 (0.116)	0.040 (0.088)	0.030 (0.101)
Moved at [14-17] \times Girl	0.220* (0.115)	0.086 (0.087)	0.112 (0.091)	0.091 (0.139)	0.027 (0.106)	0.073 (0.121)
Left-Behind at [0-6]	-0.065 (0.060)	-0.012 (0.045)	-0.028 (0.047)	-0.013 (0.104)	-0.029 (0.079)	0.036 (0.090)
Left-Behind at [0-6] \times Girl	0.213*** (0.081)	0.032 (0.061)	0.130** (0.064)	0.043 (0.095)	-0.015 (0.072)	0.054 (0.082)
Left-Behind at [7-13]	-0.024 (0.078)	0.071 (0.059)	-0.077 (0.062)	-0.025 (0.104)	-0.019 (0.079)	-0.011 (0.090)
Left-Behind at [7-13] \times Girl	0.152 (0.112)	-0.110 (0.084)	0.213** (0.088)	0.117 (0.128)	-0.031 (0.098)	0.169 (0.111)
Left-Behind at [14-17]	-0.116 (0.109)	-0.038 (0.082)	-0.048 (0.086)	0.038 (0.126)	0.014 (0.096)	0.055 (0.110)
Left-Behind at [14-17] \times Girl	0.150 (0.145)	-0.016 (0.109)	0.087 (0.115)	0.026 (0.159)	-0.109 (0.121)	0.080 (0.138)
N	1915	1915	1915	1915	1915	1915
Year of birth FE & Birth order	Yes	Yes	Yes	Yes	Yes	Yes
Mother FE	No	No	No	Yes	Yes	Yes
Number of mother FE				722	722	722

Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 3.12: Children's move and maternal migration

	Δ School gap (School progression)	
	(1)	(2)
Siblings in the baseline household in 2008		
Migrant (t-1)	-0.232*** (0.070)	
Migrant (t-1) \times Child moved with mom	0.459*** (0.151)	
Moved at [0-6]		-0.031 (0.188)
Moved at [0-6] \times Child moved with mom		0.590** (0.257)
Moved at [7-13] (t-1)		0.123 (0.091)
Moved at [7-13] (t-1) \times Child moved with mom		0.104 (0.168)
Moved at [14-18] (t-1)		-0.623*** (0.090)
Moved at [14-18] (t-1) \times Child moved with mom		0.090 (0.256)
N	5120	5120
Number of household FE	752	752
Baseline household FE	Yes	Yes
Wave year FE	Yes	Yes

Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table 3.13: Children's move and maternal migration (Cross-section sample)

	Achieved G12	Achieved G12 at 18	Achieved G12 at 19+	Achieved G12	Achieved G12 at 18	Achieved G12 G12 at 19+
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Overall effect						
Treat 1	-0.050 (0.061)	-0.071 (0.046)	0.065 (0.048)	-0.025 (0.092)	-0.039 (0.070)	0.069 (0.080)
Treat 1 \times Child moved with mom	0.161* (0.090)	0.113* (0.068)	-0.020 (0.071)	0.101 (0.114)	0.094 (0.087)	-0.026 (0.099)
Treat 2	-0.006 (0.051)	-0.034 (0.038)	0.049 (0.040)	0.016 (0.080)	-0.068 (0.061)	0.104 (0.069)
Girl	0.104*** (0.022)	0.110*** (0.017)	-0.004 (0.018)	0.094*** (0.025)	0.089*** (0.019)	0.016 (0.022)
N	1915	1915	1915	1915	1915	1915
Panel B: Heterogeneity by age at move						
Moved at [0-6]	-0.046 (0.087)	-0.085 (0.065)	0.079 (0.068)	-0.097 (0.124)	-0.080 (0.094)	0.042 (0.108)
Moved at [0-6] \times Child moved with mom	0.142 (0.115)	0.134 (0.087)	-0.079 (0.091)	0.130 (0.142)	0.158 (0.108)	-0.086 (0.124)
Moved at [7-13]	-0.095 (0.074)	-0.085 (0.056)	0.050 (0.059)	-0.174 (0.108)	-0.121 (0.082)	0.010 (0.094)
Moved at [7-13] \times Child moved with mom	0.234** (0.105)	0.113 (0.079)	0.046 (0.083)	0.234* (0.132)	0.148 (0.100)	0.053 (0.115)
Moved at [14-17]	-0.000 (0.077)	-0.046 (0.058)	0.071 (0.061)	0.105 (0.106)	0.055 (0.081)	0.086 (0.092)
Moved at [14-17] \times Child moved with mom	0.081 (0.117)	0.103 (0.088)	-0.057 (0.093)	-0.000 (0.144)	0.010 (0.109)	-0.020 (0.125)
Treat 2	-0.005 (0.051)	-0.034 (0.038)	0.048 (0.040)	-0.004 (0.081)	-0.070 (0.062)	0.083 (0.071)
Girl	0.104*** (0.022)	0.110*** (0.017)	-0.005 (0.018)	0.094*** (0.025)	0.090*** (0.019)	0.014 (0.022)
N	1915	1915	1915	1915	1915	1915
Year of birth FE & Birth order	Yes	Yes	Yes	Yes	Yes	Yes
Mother FE	No	No	No	Yes	Yes	Yes
Number of mother FE				722	722	722

Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Appendix

Table A1: Differences between treated and control groups (Cross-section sample)

	Treat 1 N = 455 (a)	Treat 2 N = 1292 (b)	Control group N = 168 (c)	Diff (c)-(a)	Diff (c)-(b)	Diff (b)-(a)
Achieved Grade 12	0.38	0.40	0.38	-0.01	-0.02	0.01
Age at Grade 12 achievement	18.89	19.09	19.54	0.65**	0.45	0.2
Achieved Grade 12 at age 18	0.16	0.16	0.13	-0.03	-0.03	0
Achieved Grade 12 with delay (age 19+)	0.18	0.18	0.20	0.02	0.02	0
Age at migration	9.59	24.01	23.82	14.22***	-0.2	14.42***
Girl	0.54	0.47	0.53	-0.01	0.06	-0.06**
Age in 2017	24.45	26.66	33.02	8.57***	6.36***	2.20***

***p<0.01, **p<0.05, *p<0.1

Table A2: Migration and School progression - Controlling for time-varying shock

	Δ School gap (School progression)	
	(1)	(2)
Panel A: all children in the baseline household in 2008		
Migrant (t-1)	-0.023 (0.029)	
Moved at [0-6]		0.333*** (0.051)
Moved at [7-13] (t-1)		0.223*** (0.035)
Moved at [14-18] (t-1)		-0.609*** (0.042)
A parent died between 2008-2017	-0.024 (0.050)	-0.021 (0.050)
N	14620	14620
Number of household FE	1884	1884
Panel B: siblings in the baseline household in 2008		
Migrant (t-1)	-0.138** (0.063)	
Moved at [0-6]		0.303** (0.122)
Moved at [7-13] (t-1)		0.142* (0.076)
Moved at [14-18] (t-1)		-0.613*** (0.085)
A parent died between 2008-2017	0.027 (0.101)	0.014 (0.100)
N	5120	5120
Number of household FE	752	752
Baseline household FE	Yes	Yes
Wave year FE	Yes	Yes

School gap is computed for children aged 7 years old and more in each wave from 2008 to 2017. Δ school gap is the difference in school gap between waves (2010-2008 ; 2012-2010 ; 2014-2012; 2017-2014). Δ school gap can be interpreted as school progression. Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table A3: Children's migration and School progression by gender - Controlling for time-varying shock

	Δ School gap (School progression)	
	(1)	(2)
Siblings in the baseline household in 2008		
Migrant (t-1)	-0.080 (0.078)	
Migrant (t-1) \times Girl	-0.114 (0.094)	
Moved at [0-6]		0.343** (0.149)
Moved at [0-6] \times Girl		-0.065 (0.216)
Moved at [7-13] (t-1)		0.182* (0.099)
Moved at [7-13] (t-1) \times Girl		-0.074 (0.125)
Moved at [14-18] (t-1)		-0.567*** (0.119)
Moved at [14-18] (t-1) \times Girl		-0.095 (0.161)
A parent died between 2008-2017	0.026 (0.100)	0.012 (0.100)
N	5120	5120
Number of household FE	752	752
Baseline household FE	Yes	Yes
Wave year FE	Yes	Yes

Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table A4: Children's migration and School progression by gender and type of move - Controlling for time-varying shock

	Δ School gap (School progression)			
	Rural-Rural move (1)	Urban-Urban move (2)	Rural-Urban move (3)	Urban-Rural move (4)
Siblings in the baseline household in 2008				
Moved at [0-6]	0.196 (0.239)	0.556** (0.264)	0.354 (0.362)	0.248 (0.463)
Moved at [0-6] \times Girl	0.032 (0.347)	-0.138 (0.335)	-0.185 (1.140)	-0.078 (0.775)
Moved at [7-13] (t-1)	0.118 (0.156)	0.235 (0.164)	1.052*** (0.345)	-0.123 (0.289)
Moved at [7-13] (t-1) \times Girl	-0.089 (0.193)	0.024 (0.217)	-0.847** (0.370)	0.538 (0.477)
Moved at [14-18] (t-1)	-0.421** (0.175)	-0.840** (0.326)	-0.597** (0.268)	-1.107*** (0.334)
Moved at [14-18] (t-1) \times Girl	-0.184 (0.263)	0.192 (0.408)	-0.362 (0.356)	0.788* (0.444)
A parent died between 2008-2017	-0.038 (0.112)	0.023 (0.125)	-0.011 (0.128)	-0.061 (0.133)
N	4067	3840	3656	3531
Number of household FE	578	576	521	516
Baseline household FE	Yes	Yes	Yes	Yes
Wave year FE	Yes	Yes	Yes	Yes

Standard errors in parentheses. Significance: ***p<0.01, **p<0.05, *p<0.1.

Table A5: Children's move and maternal migration - Controlling for time-varying shock

	Δ School gap (School progression)	
	(1)	(2)
Siblings in the baseline household in 2008		
Migrant (t-1)	-0.233*** (0.070)	
Migrant (t-1) \times Child moved with mom	0.460*** (0.151)	
Moved at [0-6]		-0.031 (0.188)
Moved at [0-6] \times Child moved with mom		0.591** (0.257)
Moved at [7-13] (t-1)		0.123 (0.091)
Moved at [7-13] (t-1) \times Child moved with mom		0.104 (0.168)
Moved at [14-18] (t-1)		-0.623*** (0.090)
Moved at [14-18] (t-1) \times Child moved with mom		0.090 (0.256)
A parent died between 2008-2017	0.030 (0.100)	0.016 (0.100)
N	5120	5120
Number of household FE	752	752
Baseline household FE	Yes	Yes
Wave year FE	Yes	Yes

Standard errors in parentheses. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A6 School inequalities in South Africa

Inequalities in terms of resource allocation and school quality in South Africa are documented in the literature. Using the 1993 South African Living Standards Survey combined with administrative data and the 1991 census, Case and Deaton (1999) investigate the relationship between educational resources and schooling outcomes just before the end of apartheid. They define education resources as the pupil-teacher ratio in primary and secondary schools and show large disparities indicating that the ratio is more than twice as big in black schools compared to white schools where the pupil-teacher ratio averages nineteen. School facilities in the communities of black children are poorly equipped: all Asians and Whites have access to primary schools with a library compared to 11% of black children. The authors find that increased pupil-teacher ratio lowers the educational attainment, the probability of being enrolled in school, and literacy and numeracy test scores only for black children. In addition, Branson and Lam (2017) emphasize that education funding assigned to white schools was more than ten times the resources allocated to black schools. Relying on the school monitoring survey in 2017-2018 and the South Africa's National Senior Certificate (NSC) administrative data, Köhler (2020) investigates the relationship between class sizes and learner outcomes at the school level taking into account the socioeconomic characteristics of public secondary schools. The author shows class sizes and schooling outcome inequalities indicating that large class sizes are concentrated in the poorest schools (from quintile 1 to quintile 3) in contrast to wealthier schools (quintile 4 and 5) characterized by smaller class sizes and higher National Senior Certificate pass rate. The main finding of the paper is that lower class sizes are positively correlated with learning outcomes only in wealthier schools in which teacher quality is better relative to the poorest schools. Furthermore, Spaull (2013) argues that the primary education system in South Africa is characterized by functional and dysfunctional schools leading to a dualistic education system. The functional schools receive only 25% of learners who perform relatively well on literacy and numeracy scores while 75% of children attend dysfunctional schools and have worse educational outcomes.

Conclusion générale

Les travaux empiriques de cette thèse ont porté sur l'analyse des interactions entre les dynamiques de population, les infrastructures et les migrations internes dans deux pays d'Afrique sub-saharienne en mobilisant des données démographiques longitudinales et de panel originales et relativement rares dans le contexte africain. Cette conclusion se donne un double objectif : d'une part, donner les implications de politiques publiques des résultats, et d'autre part, proposer des pistes pour de futures recherches.

Le premier chapitre montre que l'expansion de l'offre éducative en milieu rural réduit la fécondité précoce avant l'âge de 25 ans et conduit à une amélioration des comportements en matière de santé reproductive des femmes. Des investissements plus importants dans l'éducation de la part des décideurs publics pourraient accélérer la transition démographique en Afrique sub-saharienne et contribuer à la réalisation des Objectifs de Développement Durable (ODD) d'ici 2030. Dans un contexte où les ressources des Etats sont limitées, il est nécessaire de financer des politiques publiques visant à atteindre le processus de transition démographique pour une meilleure planification des interventions de développement. Par ailleurs, les conditions climatiques difficiles marquées par une fréquence élevée des événements climatiques extrêmes, notamment dans la zone sahélienne peuvent avoir des conséquences néfastes sur la population. Les résultats du deuxième chapitre montrent que les politiques de distribution massive de moustiquaires imprégnées ont permis de réduire la mortalité infantile liée au paludisme à Niakhar sur la dernière décennie caractérisée par des précipitations plus abondantes avec une forte variabilité interannuelle. Alors que la mortalité infantile liée au paludisme reste un problème de santé publique dans les pays de la zone sahélienne africaine, les résultats suggèrent que ces politiques pourraient atténuer les effets néfastes des chocs pluviométriques. Par conséquent, les décideurs politiques devraient continuer à soutenir les interventions anti-paludisme pour atteindre l'ODD 13 malgré la crise économique associée à la pandémie de Covid-19 qui pourrait compromettre les financements de ces interventions.

Un champ de recherche à explorer serait d'examiner les effets des politiques de construction massive d'infrastructures scolaires sur la mobilité sociale en Afrique sub-saharienne. Il s'agira de comprendre dans quelle mesure l'exposition des enfants à l'expansion de l'offre éducative améliore ou non leurs opportunités par rapport à leur parent. Par ailleurs, dans le contexte de l'Afrique de Sud, il serait intéressant d'examiner si les flux migratoires des enfants dans les zones urbaines provoquent un effet de congestion dans les écoles. Enfin, une piste de recherche à investiguer est de comprendre les raisons qui pourraient expliquer le fait que les filles soient seulement affectées par les chocs pluviométriques et non les garçons dans la zone de Niakhar.

RÉSUMÉ

L'Afrique sub-Saharienne fait partie des régions du monde qui posent les plus grands défis présents et à venir de peuplement et de développement du fait tant des dynamiques démographiques observées que des conditions climatiques des plus difficiles. En 2050, la population africaine est estimée à plus de deux milliards, ce qui représentera un quart de la population mondiale. Les travaux empiriques de cette thèse analysent les interactions entre les dynamiques de population, les infrastructures et les migrations internes dans deux pays d'Afrique sub-saharienne en mobilisant des données démographiques longitudinales et de panel originales et relativement rares dans le contexte africain.

Le premier chapitre examine l'effet causal d'une forte augmentation de la construction d'écoles entre 1998-2005 sur la fécondité, le mariage précoce et la santé reproductive dans la zone de l'observatoire de population de Niakhar située en milieu rural sénégalais. Le second étudie l'impact des chocs pluviométriques sur la mortalité infantile en milieu rural sénégalais à Niakhar sur la période 1985-2016. Enfin, le troisième chapitre examine l'effet des migrations internes en Afrique du Sud sur la scolarisation des enfants entre 2008 et 2017 en comparant la progression scolaire et le niveau d'éducation entre les enfants migrants et non-migrants.

MOTS CLÉS

Infrastructures, Dynamiques de population, Migrations internes, Education, Changement climatique, Afrique sub-Saharienne.

ABSTRACT

Sub-Saharan Africa is one of the regions of the world that faces the greatest current and future challenges in population and development due to the demographic dynamics observed and the most difficult climatic conditions. In 2050, the African population is estimated to be more than two billion and will represent a quarter of the world population. The empirical works in this thesis analyze the interactions between population dynamics, infrastructures, and internal migrations in two sub-Saharan African countries by mobilizing original longitudinal demographic data and panel data relatively rare in an African context.

The first chapter examines the causal effect of a sharp increase in school construction between 1998 and 2005 on fertility, early marriage, and reproductive health behavior in the Niakhar Health and Demographic Surveillance System (HDSS) in rural Senegal. The second one investigates the impact of rainfall shocks on infant mortality in rural Senegal, Niakhar, over the period 1985-2016. Last, the third chapter examines the effect of internal migrations in South Africa on schooling during the period 2008-2017 by comparing the school progression and educational attainment between migrant and non-migrant children.

KEYWORDS

Infrastructures, Population dynamics, Internal migrations, Education, Climate change, Sub-Saharan Africa.