UCLouvain
Economics School of Louvain

Université Paris 1 Panthéon-Sorbonne

École d'Économie de Paris

Essays on Education and Stages of Growth

Elisa Rizzo

27 November 2018

Composition du Jury:

David de la Croix (*UCLouvain*) (promoteur)

Bertrand Wigniolle (*Paris School of Economics and Paris 1*) (promoteur)

Fabio Mariani (*UCLouvain*)

Thomas Baudin (*IESEG Business School*)

Gani Aldashev (*Université Libre de Bruxelles*)

Président du Jury:

Hippolyte d'Albis (*Paris School of Economics and Paris 1*)

This thesis has been written within the European Doctorate in Economics - Erasmus Mundus (EDEEM), with the financial support by the European Commission.

To my family.
To all new adventures.

" L'uomo lasciò la locanda la mattina dopo. C'era un cielo strano, di quelli che corrono veloci, hanno fretta di tornare a casa. Soffiava vento da nord, forte, ma senza fare rumore.

All'uomo piaceva camminare. Prese la sua valigia e la sua borsa piena di carta, e si avviò lungo la strada che se ne andava, di fianco al mare.

Camminava veloce, senza voltarsi mai."

Oceano Mare, A. Baricco

Acknowledgements

This has been an incredible journey. Along the way there have been days where I was not sure of being able to arrive until the finish line, but I did it. And for this I have to thank all the incredible people I have met in these years.

Firstly, I would like to thank my supervisors: Prof. David de la Croix and Prof. Bertrand Wigniolle for their constant and patient guidance. I consider myself extremely lucky to have been able to work with two professors with such a knowledge and experience, who are also two kind and generous persons, behind their academic role. I have to say I have not been the easiest kind of PhD student. I am a free spirit with a love for mountaineering, Astrophysics and Ancient Greek poetry. Part of this thesis has been written on a train in India, part in Himalayan lodges and bakeries on my way to the Everest Base Camp. David and Bertrand had therefore a tough role: to teach me how to be a researcher while keeping me down to earth and to Economics. And for this I sincerely thank my supervisors.

I would also like to thank my bachelor thesis supervisor, Prof. Gabriella Berloffa, and my master thesis supervisor, which is also my coauthor now, Prof. Renzo Derosas. Gabriella was my first professor of Economics and she grafted onto me the passion for Development Economics, and the need for curiosity and open-mindedness that must characterize an economist. Renzo is my reference figure in Italy, a supervisor, a co-author and a friend, who taught me what it means to do things with passion. The same person who, as a historian and demographer, helped me to get off the track and look at things from different points of view.

Besides them, I would like to thank the other professors in the thesis committee: Prof. Gani Aldashev, Prof. Thomas Baudin, Prof. Fabio Mariani, and the president of the committee Prof. Hyppolyte d'Albis, not only for their insightful comments and questions to improve my thesis, but also for their encouragement. I have known all of them since the first year of my PhD. Some of them have also been my professors during my master, and my PhD as well. They have always been a source of inspiration, and help when needed.

My sincere thanks also go to the EDEEM program faculty and administrative members: Prof. Herbert Dawid, Prof. Ulrike Haake, Prof. Jean-Marc Bonnisseau, Prof. Francisco

Santana Ferra, Prof. Anna Zaharieva, Prof. Pietro Dindo, and Diana Grieswald, Catherine Germaine and Liliana de Freitas.

Special thanks go to Prof. Francois Maniquet and Prof. Frédéric Docquier, the EDEEM project coordinators in Louvain, but also incredible professors who have been a constant source of knowledge, support and encouragement, through all my PhD.

I thank the faculty members in Louvain la Neuve and Paris for the quality time shared during these years and the constant positive contribution to my activity as a researcher and as a teaching assistant, especially I thank Prof. Jean Hindricks, Prof. William Parienté, Prof. Bruno van der Linden, Prof. Luca Pensieroso, Prof Michel de Vroey, Prof. Johannes Johnen, Prof. Andrew Clark and Prof. Oliver Vanden Eynde.

I am deeply grateful to Claudine Stage, Stéphanie Pavlovitch, Marie Gilot, Anne Davister-Logist, Severine Dinjar and Thérèse Davio, for their help and guidance with really everything and for the kindness and positivity they bring everyday at IRES, making it an even more pleasant workplace.

Finally let me thank my families. Yes, families, as after this PhD I have a loving family in Italy but also a family of doctoral fellows I love in Paris and Brussels, or wherever they are and they will be.

I would like to thank my family: my mother Chiara, for her unconditional love, and my father Maurizio, for teaching me to climb any mountain I think is worth climbing, my sister Sara, for being the brightest and most resilient person I have ever known, my grandparents Vincenzo, Bruno, Elisa and Pina, my aunts and uncles, Agnese, Annalina, Giovanni, Paolo, Marco, Paolo, Michele and Christian, Diego and Ivana. You have always supported me in any possible way.

Thank you to my incredible friends in Italy: Anna Vanin, Sofia Turra, Chiara Trombini, Isabella Romoli, Susanna Aurora Basso, Laura Vendruscolo, Francesca Perini, Anna Didi Bettiol, Chiara Barea, Aurora Pietragalla, Alessia Triolo, Clelia Moroni, Cristina Borghilli, Enrico Trevisan, Giovanni Zalloni, Emanuele Giacalone. Thanks to your love I have accumulated an incredible number of Brussels Airline flights during this years.

Now the longest list. If I forget someone, it is due to the emotion of this moment. Thank you to all the friends I have interweaved my life with in these years in Paris and Brussels, especially thank you for eating all the cakes I was baking to relax from the stress of the phd life and for making me feel at home wherever I was. So thank you to Erika Pini, Andras Gregor, Sabina Lucansu, Koyel Sarkar, Claudio Telha, Manuela Braione, Ignacio Aravena,

Sinem Bas, Robert Somogyi, Lucia Granelli, Valeria Forlin, Francesca Truffa, Thu Hien Dao, Annalisa Frigo, Dalal Moosa, Juliana Mesen Vargas, Adam Levai, Fabio Blasutto, Hamzeh Arabzadeh, Baris Vardar, Bertrand Achou, Pablo Aguilar, Zainab Iftikhar, Mery Ferrando, Guzman Ourens, Keiti Kondi, Elisabetta Severi, Francois Courtoy, Charlotte de Montpellier, Eric Roca Fernandez, Leda Inga, Boris Chafwehe, Guillermo Santos, Benoit Decerf, Andreu Arenas, Bastien Chabé-Ferret, Paula Gobbi, Manuel Foster, Christoph Deuster, Lorenzo Bastaniello, Simon Schopohl, Aditi Dimri, Anna Petronevich, Moutaz Altaghlibi, Tugba Zeydanli. Thank you Riccardo Turati, Daniele Verdini, Léo Czajka and Yannik Schenk for being my official source of food, coffee, beer and laughs all along the year. Thank you to my Parisian family, Camilla Danielski, Paolo Francavilla, Mario Luca, Monica Zanardo, Sara Tafuro, Martino Pompeo, Khairul Bashar.

Contents

List of Figures

List of Tables

General Introduction

This study consists of three chapters in which the first two focus on education and the impact of public education policies on crime rates, while the last chapter explores the relationship between education and birth spacing and fertility. The first chapter presents a theoretical model where crime and education are endogenous results, considering the presence of government interventions to expand access to education and incentivise schooling. In the second chapter, I developed a model of household choice where individuals decide whether working in the legal or illegal sector as well as how much of their resources to invest in the education of their children. The model is calibrated and simulated using Indian data at district level to study the dynamics behind the occupation and education choices of the household, in order to characterize the conditions, in terms of education access, education quality and crime deterrence technology, that determine a decline of crime rates. The third chapter is an empirical study of the relationship between education and fertility in Sub-Saharan Africa, between economics and demography. In this work the latest Demographic and Health Surveys are used to explore the effect of the level of education of the mothers on the inter-birth intervals for higher-order births. The reason for focusing on inter-birth intervals, while omitting first births, is that the number of days between two births, or spacing, is a key element to describe and understand the reproductive attitudes of women and households, especially in countries where the demographic transition is not yet complete, like it is the case of Sub-Saharan Africa. That is studying intervals we are able to better understand the global relationship between education and fertility.

Even if the topic and the methods of the first two papers significantly differ from the third one, they all share the interest to better evaluate the role of education on economic growth. In particular, crime and violence as well as high fertility rates and population growth undermine economic investment. In this regard, the objective of this thesis is to contribute to an improved understanding of the diverse reasons and peculiar dynamics by which these factors might affect economic growth potential. The study is targeted on developing countries in which free access to education is only a recent achievement and where a greater effort is needed to improve the quality of the education system and teaching.

Education is one of the most important deterrents against crime. As a trigger of growth of human capital and productivity (Glomm and Ravikumar, 1992), education can indeed change the relative returns of legal versus illegal activities and reduce the incentives to switch to the illegal sector (Lochner, 2004; Naci Mocan, Billups and Overland, 2005, 2007). More human capital means that individuals expect a higher income from the legal activity and have a higher opportunity cost in joining the criminal sector (Lochner and Moretti, 2003; Machin, Marie and Vuijic, 2011). However, the increased aggregate productivity and wealth resulting from better access to education could represent an incentive for criminal activities because of the larger number of resources to be stolen, especially if there is no redistribution and income inequality remains high in the society (Ehrlich, 1973, 1975).

Following these arguments, the first chapter will be exploring the effects and dynamics at play when the government introduces a policy to increase the access to education, and if choosing the right education policy design is a crucial step to reduce crime despite the increase in aggregate resources.

In order to perform this analysis, I developed a parsimonious general equilibrium model where crime and education are endogenous results. Firstly, agents have to choose whether to educate themselves and become high skilled or remain low skilled, according to their talent endowment and the cost of education. This is the education choice. Secondly, low skilled agents have to decide whether to work legally, earning a wage set to the minimum, or become criminals, stealing resources from the honest share of the population. This is the occupation choice. Incentives to educate are given by a government intervention to reduce the cost of schooling. I consider two different types of government programs for public education. In both cases an education subsidy is provided to help to cover education costs. In one case the subsidy is funded by a progressive tax based on a tax rate and every agent, who want to go to school, pays an amount proportional to his income. In the other case the subsidy is financed by a redistribution tax such that richer agents transfer a bigger share of their earning and, after the transfer, all individuals have the same income. In each scenario I observe the results of both the education and occupation choices on the number of educated individuals and of criminals in the society. This enables me to isolate the elements which influence agents' decisions and the mechanisms behind these choices. I also compare the results of the two policies, to understand which is better designed to deter crime through education.

The results of the model show that when inequality persists and crime is more efficient, a stronger government effort is required, thus making the design of the education policy crucial. If a public education subsidy is funded by a progressive tax on income, the effect on crime

is controversial. With the education policy indeed, as access to education improves and the number of high skilled individuals increases, the aggregate levels of growth and productivity rise too. This might create an incentive for low skilled people to become criminals, especially when the newly generated wealth is not equally distributed among agents. To make sure that crime reduces, a stronger redistribution effort is needed, at the cost of an over-education inefficiency. With the redistribution policy, the higher income of the new high skilled people is almost entirely absorbed by the government to finance the subsidy for education. As a result, there are more high skilled individuals and less inequality in the distribution of income and crime drops. In this case, an over-education inefficiency might occur. That is the share of educated people in the society is higher than the optimal one to maximize the aggregate welfare of a centralized economy with a social planner. It is therefore crucial to develop a policy design such that both redistribution and efficiency are guaranteed. In this case we would be able to reduce crime without incurring in over-education.

The second chapter explores the dynamic relation between access to education, education quality and crime deterrence technology, in order to determine the conditions under which crime drops and to evaluate the role of education. Several tools are available to contrast crime rates with education being a viable deterrent (Lochner, 2004; Naci Mocan, Billups and Overland, 2005). However, as investigated in the first chapter, education policies might have conflicting effects. The effects of a better education can also favor criminal activities: in the case of an unequal distribution of income, the number of resources to be stolen from the rich increases accordingly (Ehrlich, 1973, 1975). The second chapter further explores this relationship by investigating the short- and long-term effects of education on crime.

This chapter features an overlapping generation model of household choice where crime and education are endogenous results: adults have to choose whether to work legally or switch to criminal activities as well as deciding how to allocate their resources between the education of their children and their own saving. The model explores these choices throughout three periods in agents' life: at first, agents attend school and accumulate human capital; secondly, as adults, they face the occupation choice between the legal and illegal sector and they decide how much investing in the education of their children; finally when they retire, they enjoy their savings but are exposed to crime. When only old people are victims of crime we can disentangle the main effects of the model. The equilibrium levels of human capital and crime will therefore result from the education investment and the occupational choices. The key parameters for these two choices are the productivity of the education system (ψ), which is a proxy of the quality of the school system in terms of infrastructures or teaching; the cost of education (ϵ), which represents the burden households take to pay for the education of

their children; and the crime deterrence technology (δ), which represents the ability of the society to prevent crime.

In the model there are three effects at work. The substitution between education and savings drives the education choice of the parents. If agents (adults in t) expect high crime in retirement (old in $t+1$) they use education as substitute of savings, to be less exposed to crime. The opportunity cost of human capital drives the occupation choice of the young adults. Higher is the level of human capital developed by the young adults, higher will be their loss if they decide to switch to the illegal sector. However, higher is the aggregate level of wealth and savings in the society, higher will be the incentive to join the illegal sector, this is the so called "cake size effect".

To analyze the quantitative implications of the theoretical model, I use a numerical exercise and counterfactual experiments, applied to Indian data on education, crime, production and population, gathered from different sources. The main results show that in the short-term positive education shocks, such as more access to education or higher quality of the education system, are effective as crime deterrence tools, while in the long-term they lose their influence. Moreover, technological changes in favor of crime deterrence appear to be useless, both in the short and long run.

In the third chapter, that I wrote together with Prof. Derosas, we study the relationship between female education and birth spacing, and therefore fertility, in thirty countries of Sub-Saharan Africa, with the peculiarity of making a more in-dept use of the information available.

The relationship between female education and fertility is one of the most investigated topics in the demographic as well as in the economic literature (Cochrane, 1979; Bledsoe et al., 1999). Many authors study the education-fertility dyad to define the nature of the relationship, despite the underlying complexities and the difficulties. Even assuming that the correlation between female education and fertility is not a statistical artefact, questions remain over the pathways through which education exerts its influence (Lloyd et al., 2000; Basu, 2002; Casterline et El-Zeini, 2014; Canning et al., 2015), as well as over the definition of the mechanisms through which education affects fertility (Behrman, 2015; Bongaarts, 2017).

In this paper, we analyze thirty countries of Sub-Saharan Africa to probe the effect of female education on marital fertility, after controlling for, and interacting with, a large number of other factors that also influence reproduction. More specifically, our focus is not on the number of births, but on the length of inter-birth intervals, what demographers refer to

as spacing. Spacing is a key determinant of fertility, together with age at first birth, or starting, and age at last birth, or stopping. We ask whether educated women are more likely to adopt longer inter-birth intervals, with beneficial effects on child and maternal well-being (Winikoff, 1983; Conde-Agudelo et al., 2012; Myrskylä and Barclay, 2017) and, ultimately, on the number of pregnancies.

Using data from the latest Demographic and Health Surveys (DHSs), we reconstruct the life histories of mothers and use Cox regressions to analyse inter-birth intervals in the last 5 years before the survey, for higher-order births. Although other scholars followed a similar approach, our analysis makes a more intensive use of the available information. Firstly, we use a Cox proportional hazards model with continuous-time and with time-dependent covariates, as opposed to the more usual discrete-time approach, as a more proper and precise way to deal with time to event analysis (Hosmer and Lemeshow, 1999). Secondly, we adopt a frailty model with random effects at the area unit level, to take into consideration the unobserved heterogeneity at the community level (Kravdal, 2002, 2012). Thirdly, we include in the models two covariates obtained from correspondence analysis and non-hierarchical clustering. Correspondence analysis makes it possible to identify the most relevant characteristics of the household where the woman lives in terms of region, wealth level, media exposure and house facilities. Then, by means of non-hierarchical clustering, we identify four clusters of families with similar characteristics. We use the same approach as far as the area units are concerned. In this way we deal with both the large number of variables provided by DHSs and their correlation. Finally, even if education is given once for all during the early life of the women, most probably before having children, its effect over their reproductive life changes in time, especially through the interactions with other variables. For this reason, we interact our covariate of interest with all the other covariates in the model (except countries) in order to study if there is a significant variation in the effect of education according to different conditions in the covariates.

We find that education remains a strong determinant of spacing, even after controlling for all other covariates, and that significant interactions exist, confirming the complex and variable links between education and the other factors affecting reproductive behavior. In particular, the fertility gap, that is the difference in the risk of having a new pregnancy between educated and uneducated women, increases with age, female autonomy, and the use of contraception. It decreases in the intermediate stages of the Demographic Transition or when the last-born dies. Furthermore, we find that large differences among countries remain, requiring further investigation.

Introduction Générale

La présente étude comprend trois chapitres. Les deux premiers portent sur l'éducation et l'impact des politiques d'éducation publique sur les taux de criminalité, tandis que le dernier chapitre explore la relation entre l'éducation et l'espacement des naissances et la fertilité. Le premier chapitre présente un modèle théorique où, compte tenu de la présence d'interventions gouvernementales pour élargir l'accès à l'éducation et encourager la scolarisation, la criminalité et l'éducation sont des résultats endogènes. Dans le deuxième chapitre, j'ai développé un modèle de choix du ménage où les individus décident s'ils travaillent dans le secteur légal ou illégal, ainsi que la part de leurs ressources à investir dans l'éducation de leurs enfants. Le modèle est calibré et simulé à l'aide de données indiennes au niveau des districts, ce qui me permet d'étudier la dynamique qui sous-tend les choix d'occupation et éducatifs des ménages. Ansi je peux caractériser les conditions, en termes d'accès à l'éducation, de qualité de l'éducation et de prévention de la criminalité, qui déterminent la baisse des taux de criminalité. Le troisième chapitre est une étude empirique de la relation entre éducation et fécondité en Afrique subsaharienne, entre économie et démographie. Dans ce travail, les derniéres Demographic and Health Surveys (DHSs) ont été utilisées pour explorer l'effet du niveau d'éducation des mères sur les intervalles intergénésique pour les naissances d'un ordre supérieur. La raison pour laquelle l'accent est mis sur les intervalles entre naissances, est que le nombre de jours entre deux naissances, *espacement*, est un élément clé pour décrire et comprendre les attitudes des femmes et des ménages en matière de procréation, en particulier dans les pays où la transition démographique n'est pas encore terminée, comme c'est le cas en Afrique sub-saharienne. En d'autres termes, cette étude nous permet de mieux comprendre la relation globale entre l'éducation et la fécondité.

Même si le sujet et les méthodes des deux premiers papiers diffèrent sensiblement du troisième, ils partagent tous l'intérêt de mieux évaluer le rôle de l'éducation dans la croissance économique. En particulier, la criminalité et la violence ainsi que les taux élevés de fécondité et la croissance démographique sapent l'investissement économique. A cet égard, l'objectif de cette thèse est de contribuer à une meilleure compréhension des diverses raisons et dynamiques particulières par lesquelles ces facteurs peuvent affecter le potentiel de croissance

économique. Par ailleurs, l'étude se concentre sur les pays en développement où l'accès gratuit à l'éducation n'est qu'une réalisation récente et où un effort accru est nécessaire pour améliorer la qualité du système éducatif et de l'enseignement.

L'éducation est l'un des principaux moyens de dissuasion contre la criminalité. En tant que déclencheur de la croissance du capital humain et de la productivité (Glomm et Ravikumar, 1992), l'éducation peut en effet modifier les rendements relatifs des activités légales et illégales et réduire les incitations à passer au secteur clandestin (Lochner, 2004, 2007; Naci Mocan, Billups and Overland, 2005). Plus de capital humain signifie que les individus s'attendent à un revenu plus élevé de l'activité légale. Ils ont donc un coût d'opportunité plus élevé en rejoignant le secteur criminel (Lochner et Moretti, 2003; Machin, Marie et Vuijic, 2011). Toutefois, l'augmentation de la productivité globale et de la richesse résultant d'un meilleur accès à l'éducation pourrait constituer un incitant à développer des activités criminelles. En effet, il existe des lors un plus grand nombre de ressources à voler, surtout s'il n'y a pas de redistribution et si l'inégalité des revenus reste élevée dans la société (Ehrlich, 1973, 1975). Suivant ces arguments, le premier chapitre explorera les effets et la dynamique en jeu lorsque le gouvernement introduit une politique pour accroître l'accès à l'éducation, et si le choix de la bonne conception de la politique éducative est crucial pour réduire la criminalité malgré l'augmentation des ressources globales. Pour réaliser cette analyse, j'ai développé un modèle d'équilibre général parcimonieux où la criminalité et l'éducation sont des résultats endogènes. Premièrement, les agents doivent choisir, en fonction de leur dotation en talents et du coût de l'éducation, s'ils veulent s'instruire et devenir hautement qualifiés ou rester peu qualifiés. C'est le choix de l'éducation. Ensuite, les agents peu qualifiés doivent décider s'ils doivent travailler légalement, et ainsi gagner un salaire fixé au minimum, ou devenir des criminels, et donc voler des ressources à la part honnête de la population. C'est le choix de la profession. Les incitations à l'éducation sont données par une intervention gouvernementale pour réduire le coût de la scolarisation. Je considère deux types différents de programmes gouvernementaux d'éducation publique. Dans les deux cas, une subvention à l'éducation est fournie pour aider à couvrir les frais d'éducation. Dans un cas, la subvention est financée par un impôt progressif basé sur un taux d'imposition, et chaque agent qui décide d'aller à l'école, paie un montant proportionnel à son revenu. Dans l'autre cas, la subvention est financée par un impôt de redistribution, de sorte que les agents les plus riches transfèrent une part plus importante de leurs revenus et, après le transfert, tous les individus ont le même revenu. Dans chaque scénario, j'observe les résultats des choix en matière d'éducation et de profession sur le nombre d'individus instruits et de criminels dans la société. Cela permet d'isoler les éléments qui influencent les décisions des agents et les mécanismes derrière ces choix. Je compare également les résultats des deux politiques, afin de comprendre laquelle est la mieux

conçue pour dissuader la criminalité par l'éducation. Les résultats du modèle montrent que lorsque l'inégalité persiste et que la criminalité est plus efficace, un effort gouvernemental plus important est nécessaire. Ainsi la conception de la politique éducative est cruciale. Si une subvention à l'éducation publique est financée par un impôt progressif sur le revenu, l'effet sur la criminalité est incertain. Grâce à la politique éducative, l'access à l'éducation s'améliore et le nombre des personnes hautement qualifiées augmente. Les niveaux agrégés de croissance et de productivité augmentent également. Cela pourrait inciter les personnes peu qualifiées à devenir des criminels, en particulier lorsque la richesse nouvellement créée n'est pas répartie en maniere égale entre les agents. Pour faire en sorte que la criminalité diminue, un effort de redistribution plus important est nécessaire, au prix d'une inefficacité de la suréducation. Avec la politique de redistribution, le revenu plus élevé des nouvelles personnes hautement qualifiées est presque entièrement absorbé par le gouvernement pour financer la subvention à l'éducation. Par conséquent, il y a plus de personnes hautement qualifiées et moins d'inégalité dans la répartition des revenus et la criminalité baisse. Dans ce cas, il peut y avoir une inefficacité de la suréducation. C'est-à- dire que la proportion de personnes instruites dans la société est plus élevée que la proportion optimale pour maximiser le bien-être global d'une économie centralisée avec un planificateur social. Il est donc crucial d'élaborer une politique qui garantisse, à la fois, la redistribution et l'efficacité. Dans ce cas, nous serions en mesure de réduire la criminalité sans entrainer la suréducation.

Le deuxième chapitre explore la relation dynamique entre l'accès à l'éducation, la qualité de l'éducation et la prevention de la criminalité. Il a pour but ainsi de déterminer les conditions dans lesquelles la criminalité diminue et d'évaluer le rôle de l'éducation. Plusieurs outils sont disponibles pour réduire la criminalité et, entre autres, l'éducation est un moyen de dissuasion viable (Lochner, 2004; Naci Mocan, Billups and Overland, 2005). Toutefois, comme nous l'avons vu dans le premier chapitre, les politiques éducatives peuvent avoir des effets incertains. Les effets d'une meilleure éducation peuvent également favoriser les activités criminelles: dans le cas d'une distribution inégale des revenus, le nombre de ressources à voler aux riches augmente en conséquence (Ehrlich, 1973, 1975). Le chapitre explore cette relation en étudiant les effets à court et long termes de l'éducation sur la criminalité. Il présente un modèle de choix de génération de ménages où la criminalité et l'éducation sont des résultats endogènes. D'un part, les adultes doivent choisir entre travailler légalement ou passer à des activités criminelles et d'autre part, ils doivent décider comment répartir leurs ressources entre l'éducation de leurs enfants et leur propre épargne. Le modèle explore ces choix au cours de trois périodes de la vie des agents. D'abord, les agents fréquentent l'école et accumulent du capital humain. Ensuite, à l'âge adulte, ils doivent choisir entre le secteur légal et le secteur clandestin, et ils décident du montant à investir dans l'éducation de leurs

enfants. Enfin, quand ils prennent leur retraite, ils bénéficient de leurs epargne mais sont exposés au crime. Lorsque seules les personnes âgées sont victimes de la criminalité, nous pouvons démêler les principaux effets du modèle. Les niveaux d'équilibre du capital humain et de la criminalité résulteront donc de l'investissement éducatif et des choix professionnels. Les paramètres clés de ces deux choix sont la productivité du système éducatif (ψ), qui est un indicateur de la qualité du système scolaire en termes d'infrastructures ou d'enseignement; le coût de l'éducation (ϵ), qui représente le fardeau des ménages pour payer l'éducation de leurs enfants; et la technologie de prevention (δ), qui représente la capacité de la société à prévenir la criminalité. Dans le modèle, trois effets sont à l'œuvre. La substitution entre l'éducation et l'épargne détermine le choix des parents en matière d'éducation. Si les agents (adultes en t) s'attendent à un taux de criminalité élevé à la retraite (personnes âgées en $t+1$), ils utilisent l'éducation comme substitut de l'épargne, pour être moins exposés au crime. Le coût d'opportunité du capital humain détermine le choix professionnel des jeunes adultes. Plus le niveau de capital humain développé par les jeunes adultes est élevé, plus leur perte sera grande s'ils décident de passer au secteur illégal. Toutefois, plus le niveau global de richesse et d'épargne dans la société est élevé, plus l'incitation à rejoindre le secteur clandestin sera forte, c'est ce qu'on appelle le "cake size effect". Pour analyser les implications quantitatives du modèle théorique, j'utilise un exercice numérique et des tests contrefactuels, en utilisant des données indiennes sur l'éducation, la criminalité, la production et la population, recueillies de différentes sources. Les principaux résultats montrent qu'à court terme, les chocs positifs en matière d'éducation, tels qu'un meilleur accès à l'éducation ou une meilleure qualité du système éducatif, sont efficaces en tant qu'outils de dissuasion de la criminalité, tandis qu'à long terme, ils perdent leur influence. De plus, les changements technologiques en faveur de la prevention de la criminalité semblent inutiles, tant à court qu'à long terme.

Dans le troisième chapitre, écrit avec le Prof. Derosas, nous étudions la relation entre l'éducation des femmes et les intervalles intergénésique dans trente pays de l'Afrique subsaharienne, avec la particularité de faire un usage plus approfondi des informations disponibles. La relation entre l'éducation des femmes et la fécondité est l'un des sujets les plus étudiés dans la littérature démographique et économique (Cochrane, 1979 ; Bledsoe et al., 1999). De nombreux auteurs étudient la dyade éducation-fertilité pour définir la nature de la relation, et ce, malgré les complexités et les difficultés sous-jacentes. Même en supposant que la corrélation entre l'éducation des femmes et la fécondité n'est pas un artefact statistique, les questions portent sur les voies par lesquelles l'éducation exerce son influence (Lloyd et al., 2000; Basu, 2002; Casterline et El-Zeini, 2014; Canning et al.,2015), ainsi que sur la définition des canaux par lesquels l'éducation affecte la fécondité (Behrman, 2015; Bongaarts, 2017).

Partant de là, nous analysons dans cet article la fécondité conjugale dans trente pays d'Afrique sub-saharienne pour vérifier si l'impact de l'éducation reste pertinent même après avoir pris en compte et interagi avec un grand nombre d'autres facteurs qui influencent également la fécondité. Nous ne nous concentrons pas sur l'effet de l'éducation sur le nombre de naissances, mais sur l'effet de l'éducation sur les intervalles de naissance. La raison pour laquelle nous nous intéressons aux intervalles intergénésiques est que le nombre de jours entre deux naissances, ou espacement, est un élément clé pour décrire et comprendre les attitudes reproductives des femmes et des ménages. Dans les pays où la transition démographique n'est pas encore achevée, comme c'est le cas pour l'Afrique sub-saharienne, l'espacement est un aspect fondamental déterminant de la fécondité (Cinnirella, Klemp, Weisdorf, 2012, 2012, 2017, , 2017). Nous devons donc étudier les intervalles de naissance pour mieux comprendre la relation globale entre l'éducation et la fécondité. En utilisant ces intervalles comme unités d'analyse, nous obtenons des informations très précieuses. C'est comme ouvrir la boîte noire de la relation entre l'éducation et la fécondité pour mieux la comprendre.

En utilisant les données des dernières Demographic and Health Surveys (DHSs), nous reconstruisons le cycle biologique des mères et utilisons les régressions de Cox pour analyser les intervalles entre les naissances au cours des cinq dernières années précédant l'enquête, pour les naissances d'un ordre supérieur. Bien que d'autres chercheurs aient suivi une approche similaire, notre analyse a la particularité de faire un usage plus intensif de l'information disponible. Premièrement, nous utilisons un modèle de risques proportionnels de Cox avec des covariables en temps continu et avec des covariables dépendantes du temps, par opposition à l'approche plus habituelle du temps discret, comme moyen plus approprié et plus précis de traiter l'analyse de survie (Hosmer et Lemeshow, 1999). Ensuite, nous adoptons un modèle frailty avec des effets aléatoires au niveau de zones géographiques déterminées, pour tenir compte de l'hétérogénéité non observée au niveau communautaire (Kravdal, 2002, 2012). Troisièmement, nous incluons dans les modèles deux covariables obtenues à partir de l'analyse de correspondances et du clustering non hiérarchique. L'analyse de correspondances permet d'identifier les caractéristiques les plus pertinentes du ménage où vit la femme en termes de région, de niveau de richesse, d'exposition médiatique et d'équipement. Ensuite, au moyen d'un clustering non hiérarchique, nous identifions quatre regroupements de familles ayant des caractéristiques similaires. Nous utilisons la même approche que pour les zones géographiques déterminées. Nous traitons ainsi à la fois le grand nombre de variables fournies par les DHSs, d'une part, et leur corrélation, d'autre part. Enfin, même si l'éducation est donnée une fois pour toutes au cours de l'enfance des femmes, très probablement avant d'avoir des enfants, son effet sur leur fertilité change avec le temps, en particulier à travers les interactions avec les autres variables. Pour cette raison, nous interagissons

notre covariable d'intérêt avec toutes les autres covariables du modèle (à l'exception des pays), afin d'étudier s'il y a des variations significatives de l'effet de l'éducation en fonction des différentes conditions des covariables.

Nous constatons que l'éducation reste un déterminant important de l'espacement, même après le contrôle de toutes les autres covariables, et qu'il existe des interactions significatives qui confirment les liens complexes et variables entre l'éducation et les autres facteurs affectant le comportement reproductif. En particulier, l'écart de fécondité, c'est-à-dire la différence de risque d'avoir une nouvelle grossesse entre les femmes instruites et non instruites, augmente avec l'âge, l'autonomie féminine et l'utilisation de la contraception. Il diminue, par contre, dans les étapes intermédiaires de la transition démographique, et lorsque le dernier-né meurt. En outre, nous constatons que de grandes différences subsistent entre les pays, ce qui nécessite un examen plus approfondi.

1 Can Education Reduce Crime?

A model on Public Education Policies, Inequality and Crime

Abstract

In a society characterized by an unequal distribution of income, a policy intervention to expand access to education might have contrasting effects on crime. On the one side, education deters crime by fostering the growth of human capital and productivity, thus resulting in a higher opportunity cost for individuals to switch to the illegal sector. On the other side, more human capital means also more aggregate resources to be stolen, which might increase the level of crime, if inequality persists. In this regard, the aim of this paper is to investigate the effects of a government introducing a policy designed to improve access to education. In particular, the paper explores what criteria in the design of such a policy intervention are associated with a reduction of crime despite of the raise in the aggregate wealth generated by human capital growth. In order to achieve this, I developed a parsimonious general equilibrium model where crime and education are endogenous results. Incentives to educate are given by a government intervention to expand access to education. The results of the model show that the design of the education policy is crucial. If a public education subsidy is funded with a progressive tax on income, the effect on crime is controversial. To be sure to reduce crime a stronger redistribution program is needed, at the cost of an over education inefficiency.

1.1 Introduction

Education is one of the most important practices to be used to deter crime. However, in a society characterized by inequality in the distribution of income a policy intervention to expand access to education might have conflicting effects on crime.

As a trigger of growth of human capital and productivity (Glomm and Ravikumar, 1992), education can change the relative returns of legal versus illegal activities and reduce the incentives to switch to the illegal sector (Lochner, 2004, 2007; Naci Mocan, Billups and Overland, 2005). More human capital means that individuals expect to earn a higher income from the legal activity and have a higher opportunity cost to join the criminal sector (Lochner and Moretti, 2003; Machin, Marie and Vuijic, 2011). However, the role of education as a trigger of human capital and productivity growth can also increase criminal incentives, as it increases the amount of resources to be stolen, especially if there is no redistribution and income inequality remains high in the society (Ehrlich, 1973, 1975). In this regard, the aim of this paper is to investigate the effects of a government introducing a policy designed to improve access to education. In particular, the paper explores what criteria in the design of such a policy intervention are associated with a reduction of crime despite of the raise in the aggregate wealth generated by human capital growth. In order to achieve this, I developed a parsimonious general equilibrium model where crime and education are endogenous results.

At first, the education choice implies that agents are faced with the choice of educating themselves or remaining low skilled, according to their talent endowment and the cost of education. Secondly, low skilled agents are confronted with the occupation choice as they have to choose whether staying in the legal sector, earning a wage set to the minimum, or becoming criminals, stealing resources from the honest share of the population. In this framework, the cost of education is subsidized by government incentives by different means. In one case, a progressive tax based on a tax rate results in education being funded by a contribution proportional to agents' income. In the other case the subsidy is funded by a redistribution tax such that richer agents transfer a bigger share of their earning and, after the transfer, all individuals have the same income. The creation of two scenarios allows to better observe the main drivers and influences behind the education and occupation choices before mentioned. It also favors the comparison of the two policies in order to assess the better intervention against crime. In particular, results show that in the scenario of a progressive tax the effect on crime is controversial. With the education policy, participation to education increases as also the number of high skilled in the society and aggregate wealth. This might constitute an incentive for low skilled individuals to steal those resources especially when the new wealth is not equally distributed among agents. To tackle this chance, a stronger effort by the government in terms of the implementation of a redistribution policy becomes necessary. In particular, the wealth generated by the higher income of the new high skilled individuals has to be taken by the government to pay the subsidy for education. As a result, the number of high skilled individuals further increases while resulting in less inequality in the distribution of income and ultimately in a fall of crime rates. In this case I might incur in

an over education inefficiency. That is the share of educated people in the society is higher than the optimal one to maximize the aggregate welfare of a centralized economy with a social planner. It is therefore crucial to develop a policy design such that both redistribution and efficiency are guaranteed. In this case I would be able to reduce crime without incurring in over education.

To summarize, the model features two different mechanisms: the redistribution effect and the income effect. Under the redistribution effect, the flow of resources from the people with high talent to the people with low talent helps the latter to have access to education and to improve their income legally. Higher is the level of redistribution, higher is the number of individuals who can afford education while lower is the number of low skilled who might face the occupational choice and decide to be criminals. However, under the income effect, as previously mentioned, the increased in aggregate productivity and resources deriving from the larger number of educated people in the society could represent an incentive to criminal activities, especially if income inequality persists.

The above-mentioned dynamics, with particular regard to the higher opportunity cost deriving from better access to education, are well-grounded in the theoretical and empirical literature. Naci Mocan, Billups and Overland (2005) present a dynamic model of crime in which every individual possesses two types of human capital: legal and illegal. Changing the relative returns of these capitals in favor of the legal one, the individual should switch to the legal sector easily. A way to change these relative returns is education (Lochner, 2004, 2007). There are three channels, market and non-market related, through which education can reduce crime incentives: the legal income effect, more education means more income and so higher return from the legal activity; the time availability effect, more education means more time spent in school and less time available to commit crime; and the change in patience and risk aversion, more educated agents value more the future and avoid risky criminal present gains (Witte and Tauchen, 1994; Usher, 1997; Fajnzylber et al., 2002; Lochner and Moretti, 2003; Buonanno and Leonida, 2006, 2009; Machin, Marie and Vuijic, 2011). Therefore, especially when we can observe a gap in the education attainment, between criminals and the general population, education is a powerful instrument against crime and violence (Lochner and Moretti, 2003; Harlow, 2003; Machin, Marie and Vuijic, 2011; De Baun and Roc, 2013).

Another channel through which education can help to reduce crime is income redistribution (Ehrlich, 1975; Saint-Paul and Verdier, 1992; Fernandez and Rogerson, 1998 and 1999). Financing education via a redistribution of resources from the rich to the poor is a key element for an efficient evolution of income distribution, whereas inequality may enhance crime and violence, which depress the social and economic climate, undermining economic

investment and growth potential (Ehrlich, 1973; Fernandez and Rogerson, 1998 and 1999; Bourguignon, 1999; Machin and Meghir, 2004). It is therefore vital to equalize the access to education opportunity and training. A society with equal education opportunities, will also be a society with a more equal distribution of income (Ehrlich, 1975).

Moreover, despite the abundance of literature, there is no empirical work, both micro or macro, focusing on the income effect. The reason is probably that it is a general equilibrium effect and therefore it is not identified by empirical research.

To conclude, the existing literature often addresses the key issues around inequality, education and crime, either from the point of view of the inequality-crime dyad, either from the point of view of the education-crime dyad, but never together. Therefore, I consider useful to develop a parsimonious general equilibrium model to address all the issues in the same framework. Starting from a situation of income inequality, the model simplifies the reality around two core choices, education and occupation. The intervention of the government on education acts as a redistribution channel and changes the distribution of income. The research question of the paper is therefore crucial. If we want to use education, how should we design an effective, and efficient, public school policy?

1.2 The Model

1.2.1 General Framework

Consider a static model. Individuals belong to a population of size N, normalized to 1. Each individual is born with a stock of talent h. Talent is distributed uniformly across agents over the interval $[\underline{h}; \bar{h}]$, following a uniform distribution function G(.). There are two core choices individuals have to take: the education choice (if to educate and become high skilled or not), and the occupation choice (if to be honest or criminal). The education choice concerns all agents. If they decide to educate they pay the cost of education ϵ, and they become high skilled. Their income is therefore equal to their talent endowment (h). If they decide to stay out of school they do not pay ϵ and they remain low skilled. Each low skilled receives the same minimum income, which corresponds to the minimum level of talent ($\underline{h}$). Given that ϵ is the same for every agent, no matter their level of talent, only more talented people decide to educate. The low skilled can however choose to become criminals. In this case they will earn a criminal gain which results from the total amount stolen from the honest people divided by the total amount of criminals in the society.

$q(\delta)$ represents the proportion of income stolen from each honest agent. I set $q(\delta) = q_0\sqrt{\delta}$. δ represents the share of criminals in the society and cannot be higher than the proportion of unskilled agents in the economy, given that the occupation choice concerns only the low skilled. q_0 represents the efficiency of the criminal process, and it is independent from the characteristics of the victims. Being more talented, or more educated, does not help honest agents to elude crime.

To define $q(\delta)$ I can think of a matching between criminals and honest people, analogous to a matching between unemployed workers and job vacancies. In the labor economics literature the matching function $m(u, v)$ is a homogeneous function of degree one which represents the process to create a new job, that is the rate at which a job vacancy (v) and an unemployed agent (u) meet on the job market (Pissarides, 1979, 1985; Diamond, 1982; Mortensen, 1982a, 1982b; Mortensen and Pissarides, 1994). In this framework all agents are risk neutral and I can use the homogeneity assumption to obtain $q = \frac{m(u,v)}{v}$ as the vacancy matching rate, and $p = \frac{m(u,v)}{u}$ as the rate at which unemployed workers find a job (Yashiv, 2007).

In my framework I can consider criminals ($i = \delta$) to be the job seekers and honest people ($l = 1-\delta$) to be the job vacancies. Therefore the matching function $m(i, l)$ is a homogeneous function of degree one representing the number of crimes occurring in the society, and $q = \frac{m(i,l)}{l}$ is the proportion of crimes per honest agent, or the probability of being stolen when one is honest.
However as agents are risk neutral and are all exposed to crime,

$$q = \frac{m(i,l)}{l} = m(\frac{i}{l}, 1) = m(\frac{\delta}{1-\delta}, 1) \tag{1.1}$$

can be interpreted also as the fraction of income taken by criminals per honest agent.
For a given specification of $m(x, 1)$, with $x = \frac{i}{l}$ such as

$$m(x, 1) = q_0\frac{\sqrt{x}}{\sqrt{1+x}} \tag{1.2}$$

I get, considering that $x = \frac{i}{l} = \frac{\delta}{1-\delta}$:

$$q(\delta) = q_0\sqrt{\delta} \tag{1.3}$$

Eq.1.3 is increasing in δ, the share of criminals, with diminishing returns [1].

[1] If $q(\delta) = q_0$ I can get rid of concavity and decreasing returns of the criminal activity, with some gains in tractability. However these gains are not enough to justify the change of specification of $q(\delta)$. Moreover the presence of decreasing returns may be considered as a proxy of all the crime deterrence practices

The education and the occupation choice determine the presence of three types of agents in the model: high skilled, low skilled honest and low skilled criminals.

The income of the high skilled agents is equal to their talent, net the resources stolen by criminals and the cost of education. Crime steals from the high skilled before they pay the education cost, as ϵ is like a time cost and cannot be taken by the illegal sector. An example of this cost can be the time spent in compulsory schooling to became skilled, or the time spent to acquire a mandatory qualification to perform an high skilled profession (lawyer, architect, doctor, ...). In this case agents know they will have to pay for the education cost in order to be high skilled, no matter if some part of their income is stolen, and the presence of crime does not reduce the amount of education the agent profits, once the decision to educate is taken.

$$w^{HS} = h(1 - q(\delta)) - \epsilon \tag{1.4}$$

The income of the low skilled corresponds to the minimum income, equal to the lowest level of talent possible, net the resources taken by criminals.

$$w^{LS} = \underline{h}(1 - q(\delta)) \tag{1.5}$$

The income of criminals corresponds to the criminal gain, which is the sum of all resources stolen from honest agents, divided by the share of criminals in the society.

$$w^{LSI} = \frac{q(\delta)}{\delta}(\frac{\tilde{h} - \underline{h}}{\bar{h} - \underline{h}} - \delta)\underline{h} + \frac{q(\delta)}{\delta}\int_{\tilde{h}}^{\bar{h}} \frac{h}{\bar{h} - \underline{h}} dh \tag{1.6}$$

The share of high skilled in the society is defined by $\tilde{h}$, which is the talent threshold level which belongs to the marginal individual indifferent between being high skilled or low skilled. Every individual with $h > \tilde{h}$ chooses to educate, while every individual with $h < \tilde{h}$ has not enough resources to go to school and remains low skilled. I can obtain $\tilde{h}$ comparing the income of high skilled (eq.1.4) with the income of low skilled (eq.1.5). $\tilde{h}$ is, for high skilled,

and dynamics not included in the model (like punishment measures, police enforcement or the risk of detention), as well as any social initiative against crime, which might take place when crime increases in the society. Finally if $q(\delta)$ is not concave and independent from δ, the income effect is too strong in relation to the redistribution effect and the contrast between the two mechanisms at equilibrium is less interesting in relation to the research question.

the limit income. The individual with a talent equal to $\tilde{h}$ earns as much as a low skilled agent.

The share of criminals in the society is defined by δ. I can obtain δ from the comparison between the criminal gain (eq.1.6) and the income of the low skilled honest (eq.1.5), as at equilibrium the two groups have the same income. In this case I am assuming that the occupational choice concerns only the low skilled, and not the entire population. This particular hypothesis is made to take into account the importance of income inequality as a trigger of crime, as I consider that only low skilled, that is individuals at the bottom of the distribution of talent and income, can choose to be criminal instead of honest.

According to this I can characterize the private equilibrium, where any government incentive to educate is given.

Definition 1.1: *An equilibrium with private education funding, is a vector of talent threshold and criminal share $\{\tilde{h}^*, \delta^*\}$, solution of the education choice and the occupation choice of individuals:*

- *Education Choice:*

$$\tilde{h}(1 - q(\delta)) - \epsilon = \underline{h}(1 - q(\delta)) \tag{1.7}$$

- *Occupation Choice:*

$$\frac{q(\delta)}{\delta}(\frac{\tilde{h} - \underline{h}}{\bar{h} - \underline{h}} - \delta)\underline{h} + \frac{q(\delta)}{\delta}\int_{\tilde{h}}^{\bar{h}} \frac{h}{\bar{h} - \underline{h}} dh = \underline{h}(1 - q(\delta)) \tag{1.8}$$

Before proceeding with a more extensive analysis of this equilibrium, it is useful to address the issues of optimality and social welfare. The definition of a first best result in a centralized economy can be use as a benchmark when evaluating the equilibrium levels of education and crime associated to public education policies.

1.2.2 Social Welfare and Optimality

In this section I will discuss the first best and second best problem of a centralized social planner economy. In my model the social planner maximizes the sum of the utility of the agents using an utilitarian social welfare function (SWF). However, this is not the only specification for the SWF. Speaking of SWF means addressing the issue of ranking income distribution considering social welfare. Different forms of SWF are associated to different

degrees of aversion for inequality. The concavity of the SWF is the element that determines the degree of inequality aversion. A linear SWF corresponds to zero aversion. In this case the social welfare increases if the utility or income of one agent increases, not matter his position along the distribution of income. A Rawlsian social welfare function, where the only increase in income that matters for the social welfare is the one of the poorest individual, corresponds to infinite inequality aversion. A more general form of SWF, where we can have different levels of aversion for inequality, is the iso-elastic SWF.

$$W = \frac{1}{1-\alpha}\sum_{i=1}^{N}(U_i^{1-\alpha}) \tag{1.9}$$

In this case the parameter α decides the degree of inequality aversion. If $\alpha \to 0$, the SWF collapses to the utilitarian case. If $\alpha \to \infty$, the SWF collapses to the Rawlsian case. As α increases, concavity increases and the aversion for inequality too. From now on I will use a linear social welfare function, with no aversion for inequality. The reason why I choose this form is that when there is inequality aversion it can be that a government intervention which improves access to education may be desirable even if it increases crime, if it reduces inequality. On the contrary, when aversion for inequality is zero, a change in income inequality does not affect the SWF, it affects only the education and criminal incentives, and choices, through the redistribution and income effects. In this way I reduce the channels through which I can increase the welfare of the society, enabling me to focus only on the mechanisms I am interested in.

Using an utilitarian social welfare function I am sure that aggregate welfare increases if the number of agents who decide to educate themselves increases and decreases if crime rises. This also simplifies the analysis of the equilibrium results of the following sections, where I focus only on the equilibrium values of $\tilde{h}^*$ and δ^*, without considering the full-fledged welfare function each time.

Now, coming back to the social planner maximization problem, if a social planner maximizes the sum of the utility of all agents under the resource constraint, in the society there will be no crime ($\delta = 0,\ q(\delta) = 0$) as it is a waste of resources.

$$\begin{aligned} &\max_{c(h),\tilde{h}} && \int_{\underline{h}}^{\bar{h}} u(c(h))\ dH \\ &\text{subject to} && \int_{\underline{h}}^{\bar{h}} c(h)\ dH \leq \int_{\underline{h}}^{\tilde{h}} \underline{h}\ dH + \int_{\tilde{h}}^{\bar{h}} (h-\epsilon)\ dH \end{aligned} \tag{1.10}$$

The optimal solution of this maximization problem in terms of consumption (C) is that everyone will consume the same, no matter their human capital ($c(h) = c \; \forall \; h$). Therefore, the main problem of the social planner is the maximization of the aggregate welfare with respect to $\tilde{h}_{fb}$. Then the social planner can do any transfer and easily redistribute income.

Proposition 1.1: *In a first best allocation, the social planner sets the threshold level $\tilde{h}$ at:*

$$\tilde{h}_{fb} = \underline{h} + \epsilon \tag{1.11}$$

All agents with $h > \tilde{h}_{fb}$ are sent to school, while all agents with $h < \tilde{h}_{fb}$ do not receive any education.

Proof: Maximization of aggregate welfare:

$$\max_{\tilde{h}} \quad \int_{\underline{h}}^{\tilde{h}} \underline{h} \; dH + \int_{\tilde{h}}^{\bar{h}} (h - \epsilon) \; dH \tag{1.12}$$

As the human capital of the agents follows a uniform distribution I can rewrite the total aggregate income with the following expression:

$$\underline{h}\frac{\tilde{h} - \underline{h}}{\bar{h} - \underline{h}} + \int_{\tilde{h}}^{\bar{h}} \frac{(h - \epsilon)}{\bar{h} - \underline{h}} dh \tag{1.13}$$

Taking the derivative of eq.1.13 with respect to $\tilde{h}$ equal to 0 and solving for $\tilde{h}$ I get:

$$\tilde{h}_{fb} = \underline{h} + \epsilon \tag{1.14}$$

Eq.1.14 shows that in the first best framework only individuals with enough talent to pay the education cost go to school. In any case the social planner does not need crime, or anything else, as a tool for redistribution, and resources are allocated in the most effective and efficient way.

The first best is however just a benchmark to be used to evaluate the efficiency of the equilibrium results of the model. In a more realistic framework the central planner is constrained by crime, and the first best cannot be realized. This is why it is necessary to consider a second best problem. The main constraint of this second best is the presence of crime in the society, that is $\delta > 0$. Therefore the maximization of aggregate welfare will be done for

$\tilde{h}_{sb}$ and δ, taking into account that a fraction of the population will be criminal and that a part of the aggregate welfare will be stolen. The aggregate welfare is therefore the sum of the revenue of honest people before being stolen.

$$\begin{aligned} \max_{\delta,\tilde{h}} \quad & (\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}}-\delta)\underline{h}+\int_{\tilde{h}}^{\bar{h}}\frac{(h-\epsilon)}{\bar{h}-\underline{h}}\,dh \\ \text{subject to} \quad & \frac{q(\delta)}{\delta}(\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}}-\delta)\underline{h}+\frac{q(\delta)}{\delta}\int_{\tilde{h}}^{\bar{h}}\frac{h}{\bar{h}-\underline{h}}dh=\underline{h}(1-q(\delta)) \end{aligned} \tag{1.15}$$

Lemma 1.1: *a reduction in $\tilde{h}_{sb}$ determines an increase in the level of crime δ.(Proof: appendix A)*

When crime is a constraint for the central planner a negative association arises between the education threshold and crime. This is the income effect associated to crime. If $\tilde{h}_{sb}$ is smaller, there are more high skilled and more human capital in the society. As a result, there are more resources to be stolen and crime incentives increase.

Proposition 1.2: *In a second best allocation where the level of crime is positive, the threshold level $\tilde{h}_{sb}$ is higher with respect to the first best benchmark.(Proof: appendix B)*

Proposition 1.2 shows that the presence of crime in the society reduces the resources of honest agents, as part of the legal income is stolen. As a result the threshold level for education is higher, less individuals are able to pay the cost to go to school and $\tilde{h}_{sb} > \tilde{h}_{fb}$. That is an under-education inefficiency is present. Therefore when I introduce crime as a constraint in the maximization of aggregate welfare the results I obtain are far from the first best benchmark.

1.2.3 Equilibrium

Considering the first best and second best as a benchmark, in this model crime and education are the results of the optimal choices of adult agents, and not of the decisions of a central planner.

I can use the equivalences 1.7 and 1.8 of Definition 1.1 to represent the education choice and the occupation choice in the two dimensions crime/human capital plan.

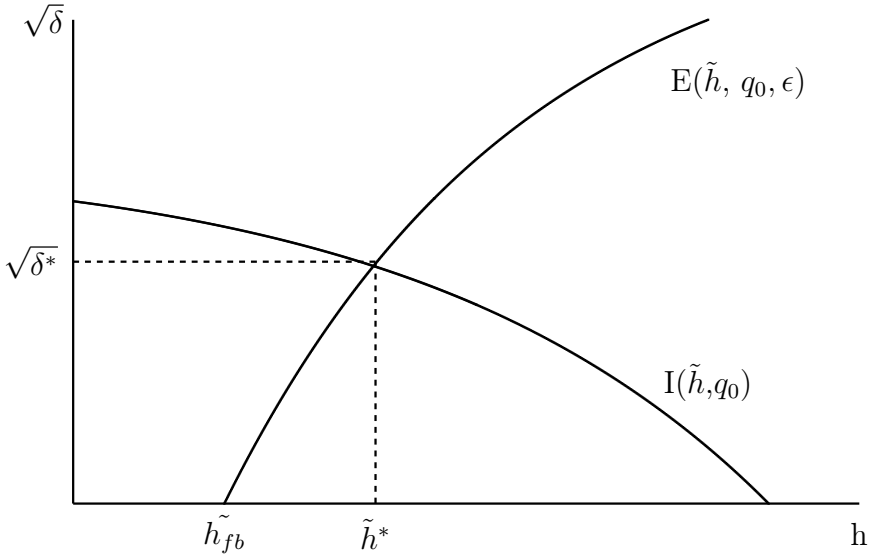
From expression 1.7 I can obtain the education choice with crime as a function of the education threshold, the efficiency of crime and the cost of education.

$$\sqrt{\delta} = \frac{\tilde{h} - \underline{h} - \epsilon}{(\tilde{h} - \underline{h})q_0} = E(\tilde{h}, q_0, \epsilon) \tag{1.16}$$

From expression 1.8 I can obtain the occupation choice with crime as a function of the education threshold and the efficiency of crime.

$$\sqrt{\delta} = (\frac{\tilde{h} - \underline{h}}{\bar{h} - \underline{h}} + \frac{\bar{h}^2 - \tilde{h}^2}{2\underline{h}(\bar{h} - \underline{h})})q_0 = I(\tilde{h}, q_0) \tag{1.17}$$

Figure 1.1: Equilibrium with Private Education Funding



Proposition 1.3: *There exists a unique equilibrium with private education funding ($\tilde{h}^*, \delta^*$), with $\tilde{h}^* \in]\underline{h}; \bar{h}[$ and $\delta^* > 0$. The education threshold at equilibrium is higher than the first best benchmark: $\tilde{h}^* > \tilde{h}_{fb}$.*

Proposition 1.4: *Comparative Statistics:*

1. *the optimal education threshold* $\tilde{h}^*$ *increases with the cost of education* ϵ*;*
2. *the optimal education threshold* $\tilde{h}^*$ *increases with the efficiency of the criminal process* q_0*.*

See appendix C for proof.

Proposition 1.3 states that when crime is a possible occupation and the cost of education is entirely paid by the agent, the amount of crime in the society is positive. The possibility to choose a criminal occupation generate therefore a first inefficiency, the crime inefficiency. As a result individuals have less resources to invest in education and the talent threshold at equilibrium is higher than the optimal one under the first best. Therefore the equilibrium corresponds to an under investment in education and human capital with respect to the first best benchmark. This is a second kind of inefficiency, the under-education inefficiency.

Proposition 1.4 shows that if the cost of education increases ($\epsilon \uparrow$) there is less interest to invest in education and become high skilled, as going to school is more expensive. In this case the education choice curve ($E(\tilde{h}, q_0, \epsilon)$) shifts toward the right. That is for any level of crime there are less high skilled individuals and $\tilde{h}^*$ increases. As a result the level of crime at equilibrium drops, as there are less resources to be stolen in the society.

If crime becomes more efficient ($q_0 \uparrow$) the education choice curve shifts to the right again, and at the same time the occupation choice curve shifts upward. When q_0 increases criminals are able to steal more from honest people, which are left with less resources to invest in education. The final effect of a change in q_0 on crime is ambiguous. If q_0 increases criminals are more efficient, however a lower share of the population is high skilled at equilibrium and there are less aggregate resources exposed to criminal activity.

1.3 Subsidizing Education against Crime: the Progressive Tax

1.3.1 General Framework

The aim of this article is to study the potentials of public education policies which extend the access to education to reduce crime. The first type of public education policy that I will

consider is a subsidy funded by a progressive tax on income. Every individual that decides to educate pays a tax proportional to his income while receiving a subsidy that covers the cost of education. The revenue of the uneducated is not subject to any taxation. τ is the education tax rate. η is the parameter that determines the education subsidy. The subsidy is the same for everyone. Income equations will be now slightly different, considering the government intervention.

The revenue of the high skilled is stolen after the tax has been paid.

$$w^{HS} = h(1-\tau)(1-q(\delta)) - \epsilon(1-\eta) \tag{1.18}$$

If low skilled decide to stay in the legal sector they receive the minimum income level, which is exposed to crime.

$$w^{LS} = \underline{h}(1-q(\delta)) \tag{1.19}$$

Criminals steal from high skilled and low skilled and divide the total criminal gain among each other.

$$w^{LSI} = \frac{q(\delta)}{\delta}\left(\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}} - \delta\right)\underline{h} + \frac{q(\delta)}{\delta}(1-\tau)\int_{\tilde{h}}^{\bar{h}} \frac{h}{\bar{h}-\underline{h}}dh \tag{1.20}$$

The government collects the tax from all the agents that decide to invest in education and go to school and uses this revenue to subsidy education.

$$\epsilon\eta \int_{\tilde{h}}^{\bar{h}} \frac{dh}{\bar{h}-\underline{h}} = \tau \int_{\tilde{h}}^{\bar{h}} \frac{h}{\bar{h}-\underline{h}}dh \tag{1.21}$$

The subsidy is the same for everyone. Therefore, for individuals with less talent endowment it is a greater contribution to the cost of schooling with respect to more talented individuals.

$$\epsilon\eta = \tau\frac{\bar{h}+\tilde{h}}{2} \tag{1.22}$$

Facing the education and occupational choices now agents consider also the presence of the tax and of the subsidy. This may help to reduce the cost of education and increase access to schooling, especially for the poorer, for which this cost is relatively higher considering

their low talent level. The education tax is used to collect resources to help more people to educate and become high skilled. Therefore, I would expect the policy to be effective and reduce crime. However, results are not so straightforward. An expansion in education means an increase in human capital that, at an aggregate level, means also more assets to be stolen by criminals. This may trigger new economic incentives for potential criminals and reduce the capacity of the public intervention.

1.3.2 Equilibrium

The presence of the tax and the subsidy changes the individuals' choices on education and occupation and affects the equilibrium values of $\tilde{h}$ and δ.

Definition 1.2: *An equilibrium with public education funded by a progressive tax, is a vector of talent threshold and criminal share $\{\tilde{h}^*_{pr}, \delta^*_{pr}\}$, solution of the education choice and the occupation choice of individuals:*

- *Education Choice:*

$$\tilde{h}(1-\tau)(1-q(\delta)) - \epsilon + \tau\frac{\bar{h}+\tilde{h}}{2} = \underline{h}(1-q(\delta)) \tag{1.23}$$

- *Occupation Choice:*

$$\frac{q(\delta)}{\delta}(\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}} - \delta)\underline{h} + \frac{q(\delta)}{\delta}(1-\tau)\int_{\tilde{h}}^{\bar{h}} \frac{h}{\bar{h}-\underline{h}}dh = \underline{h}(1-q(\delta)) \tag{1.24}$$

From eq.1.23 and eq.1.24 I can derive the education choice and the occupation choice to be plotted (eq.1.25 and eq.1.26). The presence of the tax shifts the curves of the equilibrium with private education funding. The education curve moves to the left, the crime curve moves down, and a new equilibrium is reached.

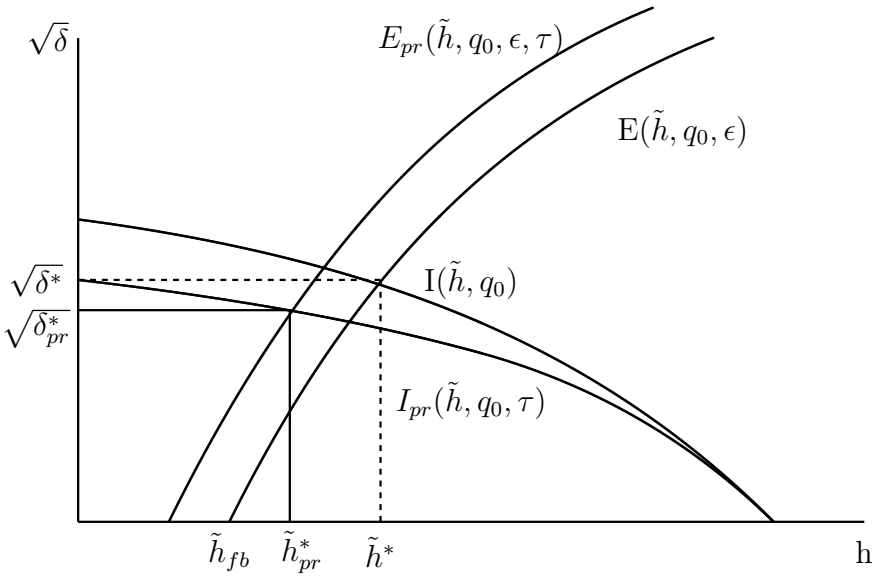
$$\sqrt{\delta} = \frac{\tilde{h}(1-\tau) - \underline{h} - \epsilon + \tau\frac{\bar{h}+\tilde{h}}{2}}{(\tilde{h}(1-\tau) - \underline{h})q_0} \equiv E_{pr}(\tilde{h}, q_0, \epsilon, \tau) \tag{1.25}$$

$$\sqrt{\delta} = ((\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}}) + (1-\tau)\frac{\bar{h}^2 - \tilde{h}^2}{2\underline{h}(\bar{h}-\underline{h})})q_0 \equiv I_{pr}(\tilde{h}, q_0, \tau) \tag{1.26}$$

$E_{pr}(\tilde{h}, q_0, \epsilon, \tau)$ is increasing in τ. The presence of the tax and of the subsidy helps to expand access to schooling.

$I_{pr}(\tilde{h}, q_0, \tau)$ is decreasing in τ, for a given $\tilde{h}$. Higher is the tax rate to be paid, lower is the amount of income that remains to be stolen.

Figure 1.2: Equilibrium with the Progressive Tax versus Equilibrium with Private Funding



Proposition 1.5: *There exists a unique equilibrium with public education funded by a progressive tax* $(\tilde{h}^*_{pr}, \delta^*_{pr})$*, with* $\tilde{h}^*_{pr} \in]\frac{h}{1-\tau}; \bar{h}[$ *and* $\delta^*_{pr} > 0$*, such that:*

1. *the education threshold decreases with* τ*,* $\tilde{h}_{fb} < \tilde{h}^*_{pr} < \tilde{h}^*$*;*
2. *the effect of* τ *on* δ^*_{pr} *is ambiguous.*

See appendix D for proof.

Proposition 1.5 shows that the introduction of the tax and the subsidy is effective to increase access to education. The optimal threshold of talent $\tilde{h}^*_{pr}$ is lower than the previous equilibrium with no public intervention. However the new threshold $\tilde{h}^*_{pr}$ is still higher than the optimal one under the centralized economy ($\tilde{h}_{fb}$). The education subsidy helps to spread education but it is not able to reproduce the first best optimal solution. This is probably

due to the fact that I am using one instrument, the education subsidy, to correct for the presence of two inefficiencies, crime and under-education. The effect of the policy on crime is ambiguous. From one end, the introduction of the tax helps more people to become high skilled and the amount of human capital in the society increases, increasing crime incentives. That is $E_{pr}(\tilde{h}, q_0, \epsilon, \tau)$ moves left for any level of crime. At the same time, as criminals steal the income of honest agents net of the tax, the tax reduces the resources exposed to criminal activities. That is $I_{pr}(\tilde{h}, q_0, \tau)$ moves downward for any level of $\tilde{h}$. The final effect on crime depends therefore on the magnitude of the shift of $E_{pr}(\tilde{h}, q_0, \epsilon, \tau)$ and $I_{pr}(\tilde{h}, q_0, \tau)$ after the introduction, or just a change, of τ.

In this regard the key parameter is the efficiency of crime (q_0). Using a simple numerical example I can show that if q_0 is high enough, crime increases rather than decreasing for low levels of τ (fig.1.3). If q_0 is small (fig.1.4) not only the equilibrium level of δ drops after the introduction of the tax, but also it decreases monotonically with τ. Therefore for high levels of crime efficiency, when the government takes action to expand education participation, the income effect dominates the redistribution effect.

Figure 1.3: Crime share evolution ($q_0 = 0.4, \bar{h} = 60, \underline{h} = 10, \epsilon = 26$)

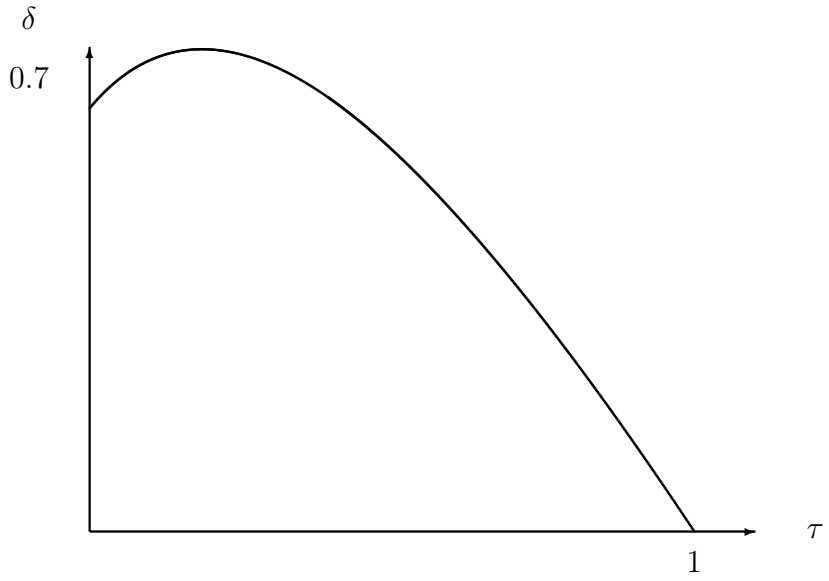
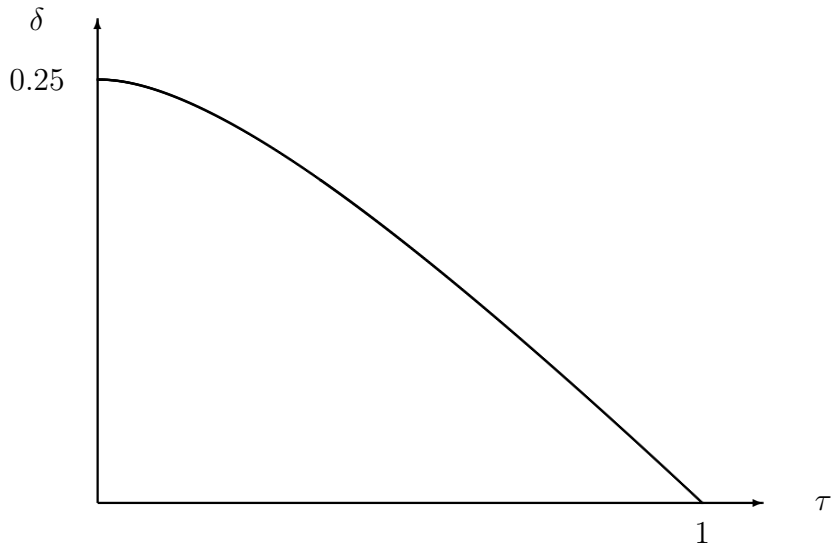


Figure 1.4: Crime share evolution ($q_0 = 0.2, \bar{h} = 60, \underline{h} = 10, \epsilon = 26$)



1.3.3 Inequality and the Progressive Tax

The progressive tax is an instrument to increase the access to education and reduce crime. However, it also affects the distribution of income in the society and so the level of inequality. In the next section I will study the results of a redistribution tax that transfers income among agents, leaving no inequality in the society. However, this is not the case with the progressive tax. It is therefore important to see the effects of the tax on the distribution of income and inequality in the society, and how this affect the two core choices of the model.

In particular, I will analyze the effect of a variation of τ on the Gini index and on the share of low skilled ($\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}}$), which are just two among other indicators of inequality. The Gini refers mostly to the inequality of the distribution of income among the high skilled, as the low skilled have all the same income (intra-group inequality). The ratio $\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}}$ indicates the share of low skilled and therefore shows us how much the society poles apart (inter-groups inequality).

The value of $\tilde{h}$ is crucial with regards to both types of inequality. Lower is $\tilde{h}$, lower is the number of low skilled and less polarized is the society as there are higher chances to be high skilled. However when $\tilde{h}$ shifts to the left along the distribution of talent, the less talented honest agent is more distant from the most talented, and there is more intra-group inequality. The opposite is true when $\tilde{h}$ is higher. Higher is $\tilde{h}$, higher is the number of low skilled earning the minimum income and more polarized is the society. However, when $\tilde{h}$ shifts to the right,

the few high skilled have all high levels of income, and the less talented honest agent is very closed to the most talented one. In this case a big share of the population earns the minimum income, while a small share of the population earns high income. The society poles apart but among the high skilled there is low inequality.

When a progressive tax is used by the government, the equilibrium level of $\tilde{h}$ is affected, as well as inter-group and intra-group inequality. Fig.1.5 and fig.1.6 show the evolution of the share of low skilled and the Gini for different values of the tax rate. For low values of τ, if q_0 is high, the public intervention is not enough to significantly increase education access and reduce income inequality. For low values of τ the share of low skilled is still high (fig.1.5). At the same time the intra group inequality increases, as even if there are new rich redistribution is low (fig.1.6). The society is therefore polarized and crime rises: the poor see themselves far from the rich and decide to become criminals to steal resources from the legal sector. Therefore, when q_0 is high and redistribution is low, education expansion may exacerbate inequality, both types, and generate the income effect which rises crime incentives. Only when the tax rate increases and, thanks to redistribution and the subsidy, more people have access to education (lower $\tilde{h}$), both inter-group and intra-group inequality drops and crime decreases as well (fig.1.3).

If τ is chosen through a majority voting process, the initial level of inequality in the society influences the attitude towards redistribution and therefore political outcomes like the decision of the tax rate (Persson and Tabellini, 1994; Saint Paul and Verdier, 1996; Benabou, 2002; Kenworthy and Pontusson, 2005). In general when there is inequality in the distribution of income, and therefore a distributional conflict, political decisions tend to produce economic policies that tax private accumulation of human capital and private investment, to promote redistributional activities. If the result of the political decision is obtained though a majority voting process, where the choice of the median voter is the winner, the position of the median voter along the income distribution is crucial (Persson and Tabellini, 1994). Normally, if the endowment of the median voter corresponds to the average endowment or is higher, the political outcome will be in favor of a non-redistributive intervention. However, when the median voter has an initial level of talent which is less than the average, he will vote for redistribution. Given that in the model low skilled agents earn just a minimum income, when the median voter is low skilled, the result of the voting process will be in favor of high redistribution (high τ), while the contrary is true if the median voter is high skilled. In this case most probably redistribution will be low (low τ) and the introduction of the education subsidy might trigger a dynamic which brings to an increase in crime.

Figure 1.5: Share of low skilled in the society ($q_0 = 0.4, \bar{h} = 60, \underline{h} = 10, \epsilon = 26$)

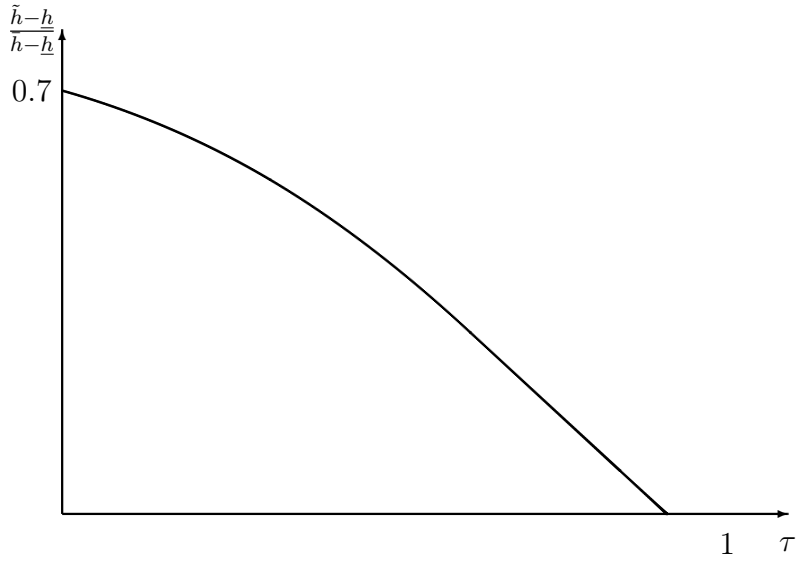
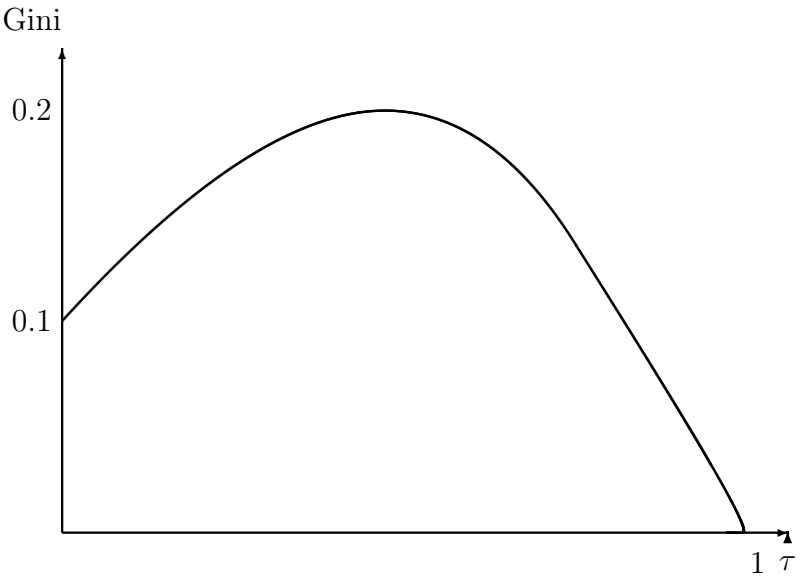


Figure 1.6: Gini Index among high skilled ($q_0 = 0.4, \bar{h} = 60, \underline{h} = 10, \epsilon = 26$)



1.4 Subsidising Education against Crime: the Redistribution Policy

1.4.1 General Framework

Even if the proportional tax helps to redistribute income in the society, some inequality remains. In the following pages I will consider a redistribution policy that totally clears inequality. This is an extreme redistribution form, however it is a useful exercise to investigate how far I can go in increasing education with a reduction in crime. I use the word extreme as in this case both educated and non educated individuals have all the same income after the policy. The income produced by the high skilled is taken by the government and redistributed to the poor as a subsidy for education.

Every high skilled agent has a human capital level equal to the education threshold $\tilde{h}$.

$$w^{HS} = \tilde{h}(1 - q(\delta)) - \epsilon(1 - \eta) \tag{1.27}$$

The income of the low skilled corresponds to the minimum income net the resources stolen by the criminals.

$$w^{LS} = \underline{h}(1 - q(\delta)) \tag{1.28}$$

The income of the criminals corresponds to the criminal gain.

$$w^{LSI} = \frac{q(\delta)}{\delta}(\frac{\tilde{h} - \underline{h}}{\bar{h} - \underline{h}} - \delta)\underline{h} + \frac{q(\delta)}{\delta}\int_{\tilde{h}}^{\bar{h}} \frac{\tilde{h}}{\bar{h} - \underline{h}} dh \tag{1.29}$$

The government collects resources from the people which decide to go to school and uses these resources to pay the subsidy.

$$\epsilon\eta \int_{\tilde{h}}^{\bar{h}} \frac{dh}{\bar{h} - \underline{h}} = \int_{\tilde{h}}^{\bar{h}} \frac{h - \tilde{h}}{\bar{h} - \underline{h}} dh \tag{1.30}$$

The subsidy is the same for everyone. Therefore for individuals with less talent endowment, it is a greater contribution to the cost of schooling with respect to more talented individuals.

$$\epsilon\eta = \frac{\bar{h} - \tilde{h}}{2} \tag{1.31}$$

Eq.1.27 shows that all educated individuals have the same revenue $\tilde{h}$. The difference between $\tilde{h}$ and their own level of talent is taken by the government to finance the education subsidy. With this mechanism a strong redistribution effect comes into play. Even if this framework is less realistic than the previous one, it gives an hint about what would happen to crime if we found a way to strongly reduce inequality of income among agents.

1.4.2 Equilibrium

The presence of the redistribution intervention changes the individuals' choices on education and occupation and affects the equilibrium values of $\tilde{h}$ and δ.

Definition 1.3: *An equilibrium with public education funded by a redistribution intervention, is a vector of talent threshold and criminal share $\{\tilde{h}^*_{re}, \delta^*_{re}\}$, solution of the education choice and the occupation choice of individuals:*

- *Education Choice:*

$$\tilde{h}(1 - q(\delta)) - \epsilon + \frac{\bar{h} - \tilde{h}}{2} = \underline{h}(1 - q(\delta)) \tag{1.32}$$

- *Occupation Choice:*

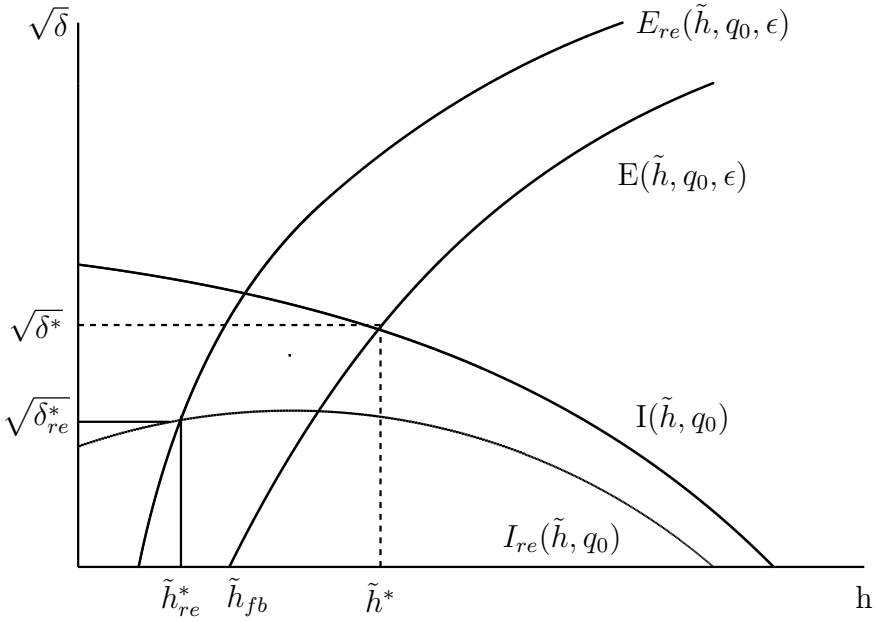
$$\frac{q(\delta)}{\delta}\left(\frac{\tilde{h} - \underline{h}}{\bar{h} - \underline{h}} - \delta\right)\underline{h} + \frac{q(\delta)}{\delta}\int_{\tilde{h}}^{\bar{h}} \frac{\tilde{h}}{\bar{h} - \underline{h}} dh = \underline{h}(1 - q(\delta)) \tag{1.33}$$

From eq.1.32 and eq.1.33 I get the education choice and the occupation choice to plot, assuming high education costs (eq.1.34 and eq.1.35). If the education cost is low then the incentives to educate are decreasing in $\tilde{h}$. If the cost of education is high the incentives to educate are increasing in $\tilde{h}$. The final effect of the redistribution policy on education and crime differs under the two cases. However I will focus only on the second one, as it is the most interesting to answer to my research question, given that when education cost is high less people have access to schooling. In the paper I assume therefore that $\epsilon > \frac{\bar{h} - \underline{h}}{2}$ and the incentive to educate is increasing in $\tilde{h}$ (See appendix E).

$$\sqrt{\delta} = \frac{\tilde{h} + \bar{h} - 2\underline{h} - 2\epsilon}{2(\tilde{h} - \underline{h})q_0} = E_{re}(\tilde{h}, q_0, \epsilon) \tag{1.34}$$

$$\sqrt{\delta} = (\frac{\tilde{h} - \underline{h}}{\bar{h} - \underline{h}} + \frac{\tilde{h}}{\underline{h}}\frac{\bar{h} - \tilde{h}}{\bar{h} - \underline{h}})q_0 = I_{re}(\tilde{h}, q_0) \tag{1.35}$$

Figure 1.7: Equilibrium with the Redistribution Tax versus Equilibrium with Private Funding



Proposition 1.6: *There exists an equilibrium with public education funded by a redistribution intervention* $(\tilde{h}^*_{re}, \delta^*_{re})$*, with* $\tilde{h}^*_{re} \in]\underline{h}; \bar{h}[$ *and* $\delta^*_{re} > 0$*, such that:*

1. *it may be that the education threshold is smaller than the first best benchmark,* $\tilde{h}^*_{re} < \tilde{h}_{fb}$*;*
2. *considering a reasonable range for the main parameters, crime drops at equilibrium,* $\delta^*_{re} < \delta^*$*.*

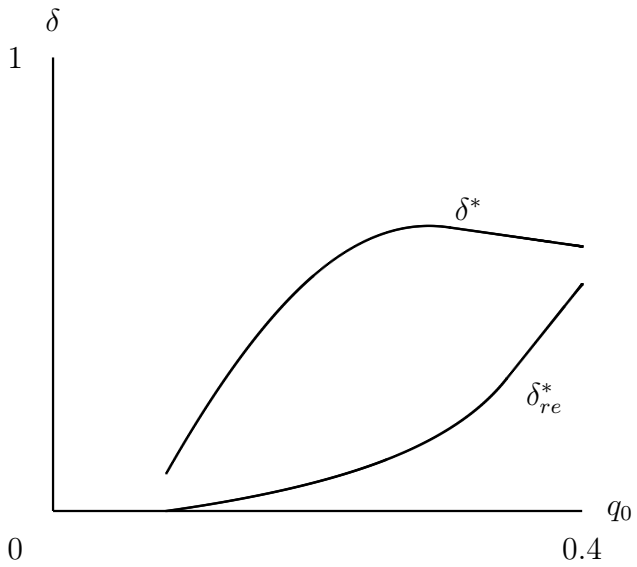
See appendix F for proof.

Proposition 1.6 shows that the redistribution policy is able to strongly increase access to education, up to the point where an over education inefficiency is originated. That is the share of educated people in the society might be higher than the optimal one to maximize the total welfare in a centralized economy. It also shows that when it is possible to significantly increase the share of high skilled in the society and, at the same time, to strongly reduce inequality, some criminal activity persists. However, the level of crime at equilibrium is lower than the equilibrium with private funding, considering a reasonable range for the

parameters (fig.1.8). In the redistribution framework everyone has the same income, there is no inequality and wealth is used to increase access to education. What is left for criminals is sufficiently low to discourage an illegal occupation.

To be more precise I can look at a numerical example. Looking at the evolution of the equilibrium value of δ without and with the redistribution policy, I can conclude that, for a range of value of q_0 up to 0.4, which is the maximum level of crime efficiency considered previously, the government intervention targeting education expansion is effective against crime. That is crime drops at equilibrium with respect to the private education funding framework.

Figure 1.8: Equilibrium value of crime in the private funding framework and in the redistribution framework, according to different values of crime efficiency (q_0). ($\bar{h} = 60, \underline{h} = 10, \epsilon = 26$)



1.4.3 The Redistribution Policy vs The Progressive Tax

Both these types of public education policy affect the number of criminals and of educated people at equilibrium, through different mechanisms. In particular, a government intervention that improves access to education can positively influence unskilled people to choose a career in the legal sector, substituting crime with education while escaping poverty (redistribution effect). Under the redistribution effect, the resources that flow from the high

talented to the low talented are used by the poor to educate themselves and to increase their income legally. Higher is the level of redistribution higher is the number of individuals that can afford education. This effect reduces the number of low skilled individuals who might face the occupational choice and switch to the illegal sector. On the other side, the increased level of aggregate wealth can represent an incentive to criminal activities (income effect). The income effect considers the consequences of an expansion of schooling on aggregate income. Under the income effect, as the number of educated people increases, the aggregate production and wealth also increase. This expansion might be appealing for criminals and potential criminals in terms of illegal gain, as more resources are now exposed to criminal activity.

Each policy intervention has its own peculiar impact on the redistribution and income effects. The main difference among the two policies is on inequality. While the progressive tax allows the new educated individuals to retain some of their increased income, under the redistribution policy every high skilled individual will have the same income after tax. The transfer of wealth among educated people here is stronger as there is no income inequality. With the redistribution policy any increase in aggregate wealth is taken by the government and used to finance education for the poor. For this reason, almost nothing is left to be stolen. The redistribution effect wins, and with the risk of the over education inefficiency, crime drops. With the progressive tax policy, the high skilled keep part of their new income resulting in more individual and aggregate resources exposed to criminal activities. For this reason under the proportional tax it might happen that the income effect wins and crime increases, especially when the efficiency of crime is high (high q_0) and the tax rate is low (low τ).

It is also important to underline that both policies can generate inefficiencies. The progressive tax helps to expand education, however it may increase crime. The redistribution policy pushes redistribution to the extreme, however the aggregate welfare might be sub-optimal with respect to the first best benchmark. It is therefore crucial to develop a policy design such that both redistribution and efficiency are guaranteed. In this case it would be possible to reduce crime without incurring in over education.

1.5 Conclusions

In a society characterized by inequality in the distribution of income, education might be a powerful instrument to reduce crime. Fighting crime with education can be done using different public schooling interventions: a subsidy financed by a progressive tax or a subsidy financed by a redistribution policy. The results of the model show that the design of the education policy is crucial. When a policy is introduced two effects are generated: the redistribution effect (flow of resources from the high talented to the low talented and decrease in income inequality) and the income effect (more schooling means more production and also more aggregate resources to be stolen). If a public education subsidy is funded by a progressive tax on income, even if more people have access to education, the effect on crime is controversial. As the new wealth of the new high skilled is not entirely taken by the government but is exposed to criminal activities, the income effect might win. There is the possibility that crime increases at equilibrium. To be sure to reduce crime a stronger redistribution intervention is needed, at the cost of an over education inefficiency. When the public education subsidy is funded by a redistribution intervention, as the new income of the new high skilled is taken by the government to increase access to education, there are not enough aggregate resources to be stolen. The redistribution effect wins and crime drops. This may come with the price of over education, which is inefficient in terms of the aggregate welfare maximization. It is therefore crucial to develop a policy design such that both redistribution and efficiency are guaranteed. In this case it would be possible to reduce crime without incurring in over education.

1.6 Appendix

Appendix A

Proof of lemma 1.1:

The main constraint of the maximization of aggregate welfare in the second best is the presence of crime, represented by the occupation choice of the low skilled individuals.

$$\frac{q(\delta)}{\delta}\left(\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}}-\delta\right)\underline{h}+\frac{q(\delta)}{\delta}\int_{\tilde{h}}^{\bar{h}}\frac{h}{\bar{h}-\underline{h}}dh=\underline{h}(1-q(\delta)) \tag{1.36}$$

Eq.1.36 can be rewritten to obtain an expression of crime as a function of $\tilde{h}$.

$$\begin{aligned}
&\frac{q(\delta)}{\delta}(\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}}-\delta)\underline{h}+\frac{q(\delta)}{\delta}\int_{\tilde{h}}^{\bar{h}}\frac{h}{\bar{h}-\underline{h}}dh=\underline{h}(1-q(\delta))\\
&\frac{q(\delta)}{\delta}((\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}}-\delta)\underline{h}+\frac{1}{\bar{h}-\underline{h}}(\frac{\bar{h}^2-\tilde{h}^2}{2}))=\underline{h}(1-q(\delta))\\
&(\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}}-\delta)\underline{h}+\frac{1}{\bar{h}-\underline{h}}(\frac{\bar{h}^2-\tilde{h}^2}{2})=\underline{h}(1-q(\delta))\frac{\delta}{q(\delta)}\\
&\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}}\underline{h}+\frac{1}{\bar{h}-\underline{h}}(\frac{\bar{h}^2-\tilde{h}^2}{2})=\frac{\underline{h}\delta}{q(\delta)}
\end{aligned} \tag{1.37}$$

Setting $I(\tilde{h})=\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}}\underline{h}+\frac{1}{\bar{h}-\underline{h}}(\frac{\bar{h}^2-\tilde{h}^2}{2})$ and taking the derivative with respect to $\tilde{h}$ I get:

$$I'_{\tilde{h}}=-\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}} \tag{1.38}$$

Therefore the expression $I(\tilde{h})$ is decreasing in $\tilde{h}$. That is whenever there is education expansion and $\tilde{h}$ is smaller, $I(\tilde{h})$ increases. From eq.1.37 I can write:

$$I(\tilde{h})=\frac{\underline{h}\delta}{q(\delta)}=\frac{\underline{h}\sqrt{\delta}}{q_0} \tag{1.39}$$

As $I(\tilde{h})$ and δ are positively related, whenever there is education expansion and $\tilde{h}$ is smaller, δ increases.

Appendix B

Proof of proposition 1.2: Eq.1.39 is another way of writing the constraint expressed in eq.1.36. Therefore the maximization problem in eq.1.15 can be rewritten as:

$$\begin{aligned}
&\max_{\delta,\tilde{h}} && (\frac{\tilde{h}-\underline{h}}{\bar{h}-\underline{h}}-\delta)\underline{h}+\frac{1}{\bar{h}-\underline{h}}(\frac{\bar{h}^2-\tilde{h}^2}{2}-\epsilon(\bar{h}-\tilde{h}))\\
&\text{subject to} && I(\tilde{h})=\frac{\underline{h}\delta}{q(\delta)}=\frac{\underline{h}\sqrt{\delta}}{q_0}
\end{aligned} \tag{1.40}$$

From the constraint I can get an expression of δ to be plugged in the welfare maximization.

$$\delta = (\frac{I(\tilde{h})q_0}{\underline{h}})^2 = \Delta(\tilde{h}) \tag{1.41}$$

$\Delta(\tilde{h})$ is decreasing in $\tilde{h}$ as $I(\tilde{h})$ is decreasing in $\tilde{h}$.
Plugging $\delta = \Delta(\tilde{h})$ in the expression for the aggregate welfare, I obtain:

$$(\frac{\tilde{h} - \underline{h}}{\bar{h} - \underline{h}} - \Delta(\tilde{h}))\underline{h} + \frac{1}{\bar{h} - \underline{h}}(\frac{\bar{h}^2 - \tilde{h}^2}{2} - \epsilon(\bar{h} - \tilde{h})) \tag{1.42}$$

Taking the derivative of eq.1.42 with respect to $\tilde{h}$ and setting the expression equal to 0, I get:

$$\underline{h}(\frac{1}{\bar{h} - \underline{h}} - \Delta^{'}(\tilde{h})) + \frac{1}{\bar{h} - \underline{h}}(\epsilon - \tilde{h}) = 0 \tag{1.43}$$

Therefore

$$\tilde{h}_{sb} = \underline{h} + \epsilon - \underline{h}(\bar{h} - \underline{h})\Delta^{'}(\tilde{h}) \tag{1.44}$$

In eq.1.44 $\underline{h} + \epsilon = \tilde{h}_{fb}$ and $\underline{h}(\bar{h} - \underline{h})\Delta^{'}(\tilde{h}) < 0$ as $\Delta(\tilde{h})$ is decreasing in $\tilde{h}$. Therefore $\tilde{h}_{sb}$ corresponds to $\tilde{h}_{fb}$ minus a negative quantity, which is like adding a positive quantity. As a result $\tilde{h}_{sb} > \tilde{h}_{fb}$.

Appendix C

Proof of proposition 1.3:
The equilibrium with private education funding exists as

- at $\tilde{h} = \underline{h}$ $I(\underline{h}, q_0) > E(\underline{h}, q_0, \epsilon)$;
- at $\tilde{h} = \bar{h}$ $E(\bar{h}, q_0, \epsilon) > I(\bar{h}, q_0)$;

so the two functions cross each other.
The equilibrium with private education funding is unique as

- $E(\tilde{h}, q_0, \epsilon)$ is increasing in $\tilde{h}$ $(\partial E/\partial \tilde{h} > 0)$;
- $I(\tilde{h}, q_0)$ is decreasing in $\tilde{h}$ $(\partial I/\partial \tilde{h} < 0)$.

Solving $E(\tilde{h}, q_0, \epsilon) = 0$ (eq.1.16) I find the following expression for $\tilde{h}$:

$$\tilde{h} = \underline{h} + \frac{\epsilon}{1 - q_0\sqrt{\delta}}$$

As $\tilde{h}_{fb} = \underline{h} + \epsilon$ and ϵ divided by a positive quantity lower than 1 is higher than ϵ then $\tilde{h} > \tilde{h}_{fb}$.

Proof of proposition 1.4:

1. education incentive is decreasing in the education cost as $\partial E/\partial \epsilon < 0$;
2. education incentive is decreasing in the efficiency of crime as $\partial E/\partial q_0 < 0$.

Appendix D

Proof of proposition 1.5:
The education choice given in eq.1.23 is satisfied only for values of $\tilde{h} > \frac{\underline{h}}{1-\tau}$, as $-\epsilon + \tau\frac{\bar{h}+\tilde{h}}{2}$ is negative and to assure that the left-end side of eq.1.23 equals the right-end side, $\tilde{h}$ cannot be too small. In particular $E_{pr}(\tilde{h}, q_0, \epsilon, \tau)$ presents an asymptote at the point $\tilde{h} = \frac{\underline{h}}{1-\tau}$.
The equilibrium with the proportional tax exists as

- $I_{pr}(\tilde{h}, q_0, \tau) > E_{pr}(\tilde{h}, q_0, \epsilon, \tau)$ at $\tilde{h}$
  equal to a value such that $E_{pr}(\tilde{h}, q_0, \epsilon, \tau) = 0$;
- $E_{pr}(\bar{h}, q_0, \epsilon, \tau) > I_{pr}(\bar{h}, q_0, \tau)$ at $\tilde{h} = \bar{h}$;

so the two functions cross each other on the right of the asymptote.

The equilibrium with the proportional tax is unique as

- $E_{pr}(\tilde{h}, q_0, \epsilon, \tau)$ is increasing in $\tilde{h}$ ($\partial E_{pr}/\partial \tilde{h} > 0$ for low values of τ);
- $I_{pr}(\tilde{h}, q_0, \tau)$ is decreasing in $\tilde{h}$ ($\partial I_{pr}/\partial \tilde{h} < 0$ for $\tilde{h} > \frac{\underline{h}}{1-\tau}$).

1. To prove that $\tilde{h}$ decreases I start from the equilibrium equality

   $E_{pr}(\tilde{h}, q_0, \epsilon, \tau) = I_{pr}(\tilde{h}, q_0, \tau)$.

   The implicit function theorem states that

$$\frac{\partial E_{pr}}{\partial \tilde{h}}\tilde{h}'(\tau) + \frac{\partial E_{pr}}{\partial \tau} = \frac{\partial I_{pr}}{\partial \tilde{h}}\tilde{h}'(\tau) + \frac{\partial I_{pr}}{\partial \tau}$$

therefore

$$\tilde{h}'(\tau) = \frac{\frac{\partial I_{pr}}{\partial \tau} - \frac{\partial E_{pr}}{\partial \tau}}{\frac{\partial E_{pr}}{\partial \tilde{h}} - \frac{\partial I_{pr}}{\partial \tilde{h}}}$$

As $\frac{\partial I_{pr}}{\partial \tau} < 0$ and $\frac{\partial E_{pr}}{\partial \tau} > 0$ the numerator is negative. As $\frac{\partial E_{pr}}{\partial \tilde{h}} > 0$ and $\frac{\partial I_{pr}}{\partial \tilde{h}} < 0$ the denominator is positive. As a result $\tilde{h}$ is decreasing in τ and there is some education expansion with the introduction of the tax.

2. The final result of τ on crime is ambiguous as τ has two contrasting effects on the education choice (eq.1.23). $\tilde{h}$ is decreasing in τ, however $E_{pr}(\tilde{h}, q_0, \epsilon, \tau)$ is increasing both in τ and $\tilde{h}$.

Appendix E

The derivative for $\tilde{h}$ of the RHS of expression (1.34) is

$$\frac{\partial E_{re}}{\partial \tilde{h}} = \frac{2\epsilon + \underline{h} - \bar{h}}{2q_0(\tilde{h} - \underline{h})^2} \tag{1.45}$$

The sign of this equation depends on the education cost "ϵ".

If the education cost is low ($\epsilon < \frac{\bar{h}-\underline{h}}{2}$) then the incentives to educate are decreasing in $\tilde{h}$. If the cost of education is high ($\epsilon > \frac{\bar{h}-\underline{h}}{2}$) the incentives to educate are increasing in $\tilde{h}$. The final effect of the redistribution policy on education and crime differs under the two cases. However I will focus only on the second one, as it is the most interesting to answer my research question. In the paper I assume therefore that $\epsilon > \frac{\bar{h}-\underline{h}}{2}$ and the incentive to educate is increasing in $\tilde{h}$.

Appendix F

Proof of proposition 1.6:
The equilibrium with private education funding exists as

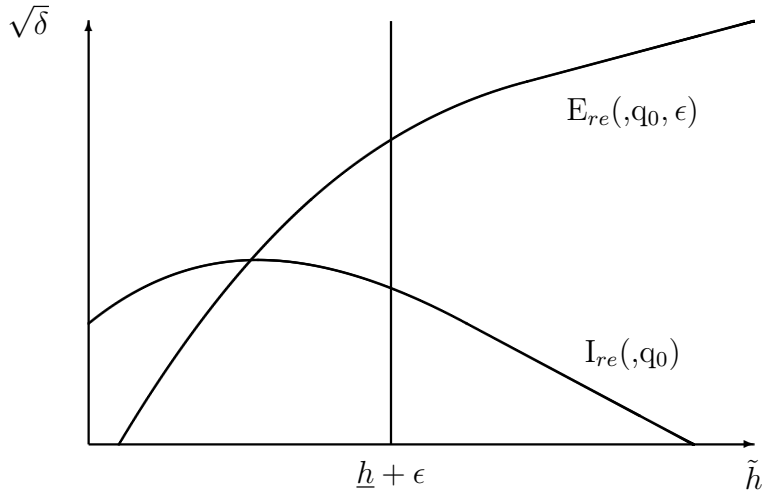
- at $\tilde{h} = \underline{h}$ $I_{re}(\underline{h}, q_0) > E_{re}(\underline{h}, q_0, \epsilon)$;
- at $\tilde{h} = \bar{h}$ $E_{re}(\bar{h}, q_0, \epsilon) > I_{re}(\bar{h}, q_0)$;

so the two functions cross each other.

The uniqueness of the equilibrium is not necessarily granted as crime incentive is first increasing and then decreasing in $\tilde{h}$. However simulations using realistic values for the parameters show that the equilibrium is unique.

1. To prove that $\tilde{h}_{re} < \tilde{h}_{fb}$ and so there is over-education, the crime choice and the education choice under the redistribution intervention should be evaluated at $\tilde{h} = \tilde{h}_{fb} = \underline{h} + \epsilon$. If the education incentive is higher than the criminal one for this value of $\tilde{h}$ this means that the two curves intersect before, as $E_{re}(\tilde{h}, q_0, \epsilon)$ is increasing in $\tilde{h}$, and so $\tilde{h}_{re} < \tilde{h}_{fb}$ (fig.1.9).

Figure 1.9: Education and Occupation choice for different values of $\tilde{h}$



The education cost parameter ϵ determines for which conditions
$E_{re}(\tilde{h}, q_0, \epsilon) > I_{re}(\tilde{h}, q_0)$ at $\tilde{h} = \underline{h} + \epsilon$.

When ϵ reaches its maximum possible value ($\epsilon = \bar{h} - \underline{h}$) then, at $\tilde{h} = \underline{h} + \epsilon$,
$E_{re}(\tilde{h}, q_0, \epsilon) = 0$ and $I_{re}(\tilde{h}, q_0) = q_0$ so $E_{re}(\tilde{h}, q_0, \epsilon) < I_{re}(\tilde{h}, q_0)$ (fig.1.10). Therefore for high values of the education cost there is no over-education as the $E_{re}(\tilde{h}, q_0, \epsilon)$ curve cannot be higher than the $I_{re}(\tilde{h}, q_0)$ curve at $\tilde{h} = \tilde{h}_{fb}$.

When ϵ reaches its minimum possible value ($\epsilon = \frac{\bar{h}-\underline{h}}{2}$), or it is close to the minimum value, then at $\tilde{h} = \underline{h} + \epsilon$:

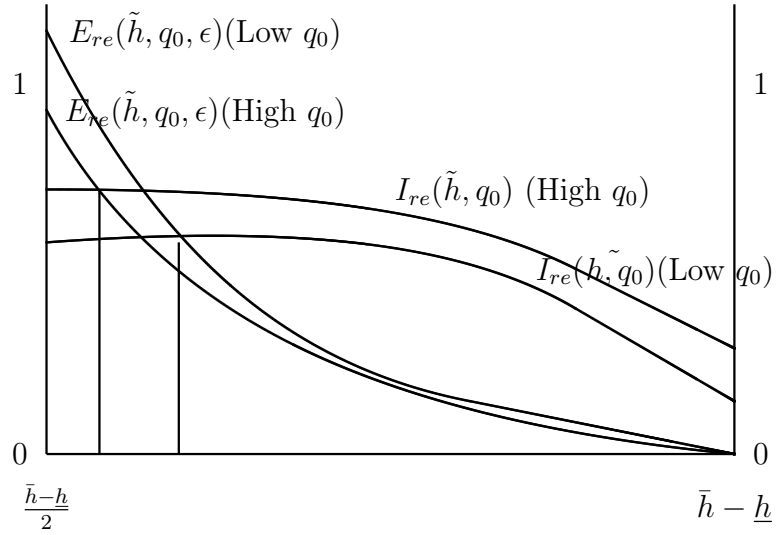
$$E_{re}(\tilde{h}, q_0, \epsilon) = \frac{1}{2q_0} \tag{1.46}$$

$$I_{re}(\tilde{h}, q_0) = (\frac{1}{2} + \frac{\bar{h} + \underline{h}}{4\underline{h}})q_0 \tag{1.47}$$

Therefore there will be over education if $\frac{1}{2q_0} > (\frac{1}{2} + \frac{\bar{h}+\underline{h}}{4\underline{h}})q_0$. This is true if q_0 is small enough and if the range of talent distribution is not too big. In this case $\frac{1}{2q_0}$ is easily

higher than $(\frac{1}{2}+\frac{\bar{h}+\underline{h}}{4\underline{h}})q_0$ and therefore the range of value of ϵ for which the introduction of the redistribution tax results in over-education is wider. This is shown in fig.1.10, where the education and criminal incentives are represented at $\tilde{h} = h_{fb}$, for different values of q_0, along the distribution of ϵ. When $E_{re}(\tilde{h}, q_0, \epsilon)$ is over $I_{re}(\tilde{h}, q_0)$, then there is over education after the redistribution intervention.

Figure 1.10: Education Incentives and Crime Incentives at $\tilde{h} = h_{fb}$, for different values of ϵ.



2 Education and Crime in India: a District Level Analysis

Abstract

Property crime imposes a cost on economies, especially in developing ones. In this paper, I study the dynamic relation between education access, education quality and crime deterrence technology to determine under which conditions crime drops and to evaluate the role of education. To perform this analysis I designed a model of household choice in which both crime and education are endogenous factors at equilibrium: adults decide whether to work legally or illegally as well as the number of resources they devote to the education of their children, given education accessibility, education quality and crime deterrence. The model is then calibrated and simulated using Indian data at the district level. As a result of this modelization strategy, positive shifts of education accessibility or quality are effective deterrents of crime in the short run, while in the long run general equilibrium effects make them ineffective. Moreover, technological changes in favor of crime deterrence appear to be useless, both in the short and long run.

2.1 Introduction

Crime represents a huge cost for economies: it hinders growth and acts like a tax on what is legally produced. In his Nobel talk of 1993, Gary Becker defined himself "puzzled" by the idea that crime is harmful for societies as it appears as an alternative way of resource redistribution from the rich to the poor. However, he concluded that crime is harmful as the resources that are stolen are not used to create wealth but for criminal activities. More recently Powell, Manish and Nair (2010) explained that crime, together with corruption, increases economic uncertainty and threatens property rights and the rule of law, undermining long term growth. Detotto and Otranto (2010) claim that crime behaves like a tax because

it depresses investments, reduces the ability of firms to compete and reallocates economic resources inefficiently. In general, despite of the size of the illegal sector, crime puts at risk the overall economic performance and growth potential of a society (Mauro and Carmeci, 2006; Detotto and Otranto, 2010; Kumar, 2013; Goulas and Zervoyianni, 2013).

Amongst the many options to prevent crime, education is an important tool (Lochner, 2004; Naci Mocan, Billups and Overland, 2005) as it increases the returns of legal activities against illegal ones. However, the effects of education on crime deterrence are ambiguous: having more educated people can in fact positively impact on crime because, when there is an unequal distribution of income, it increases the number of resources to be stolen from the rich (Ehrlich, 1973, 1975). For this reason, in this paper, I investigate if an increase in education necessarily deters crime.

To study this issue I designed an overlapping generation model of household choice where crime and education are endogenous results. Adult agents choose in which sector to operate, legal or illegal, and how to allocate their resources between the education of their children and the savings for their retirement. The model has three periods. In the first period of their lives, agents go to school and accumulate human capital. When they are adult, they face the occupation choice between the legal and illegal sector and have to decide how much investing in the education of their children. If they decide to be criminal they lose all the human capital and therefore the potential income, accumulated when in school. This assumption is justified by the type of crime I am considering: in the model I refer to blue-collar crime that does not require specific skills, therefore human capital does not increase the returns of the illegal activity. Moreover, this assumption is a proxy for the dynamics taking place in reality. Some authors claim the negative consequences of prison sentences and jail time on the labor market opportunities of ex-offenders. Decker et al. (2015) show that prior prison sentences remain important impediments to find a job for ex-offenders. As a result, they may come back to commit crime, without being able to integrate again in the society. Also Western, Kling and Weiman (2001) support the view that prison time and incarceration have large effects on the inequality of the distribution of income and potential earnings, increasing the former and decreasing the latter. To continue, in the third period, when agents retire, they enjoy their savings but are exposed to crime, no matter their occupation choice when young. When only old people are victims of crime I can disentangle the main effects of the model.

The equilibrium levels of human capital and crime will therefore result from the education investment and the occupational choices of adult agents. The key parameters for these two choices are the productivity of the education system(ψ), which is a proxy of the quality of the school system in terms of infrastructures or teaching; the cost of education (ϵ), which

represents the burden households have to take to pay for the education of their children; and the crime deterrence technology (δ), which represents the ability of the society to prevent crime.

In the model there are three effects at work. The substitution between education and savings drives the education choice of the parents. If agents (adults in t) expect high crime in retirement (old in $t+1$) they use education as substitute of savings, to be less exposed to crime. The opportunity cost of human capital drives the occupation choice of the young adults. Higher is the level of human capital developed by the young adults, higher will be their loss if they decide to switch to the illegal sector. However, higher is the aggregate level of wealth and savings in the society, higher will be the incentive to join the illegal sector, this is the so called "cake size effect".

To analyse the quantitative implications of the theoretical model I use a numerical exercise and counterfactual experiments, using Indian data on education, crime, production and population, gathered from different sources. The main results show that in the short run positive education shocks, that is more access to - or higher quality of - the education service, are effective as a crime deterrence practice, while losing their power in the long run. In addition, technological changes in favor of crime deterrence appear to be useless, both in the short and long run. The reason why the data for the numerical exercise are taken from Indian regions lies in the significant changes that have been occurring in the last years in the Indian Education System. In particular, The Right of Children to Free and Compulsory Education Act that came into force in April 2010, imposed new rules for free and compulsory education from class 1 to 8, that is for all children between 6 and 14 years. The action of the government also included a stronger effort in collecting data on schools. As a result, India has now very interesting data on education, available within the DISE project (District for Information System on Education).

In addition to education, I also use data on crime, production and population. The number of crime acts is registered in the National Crime Records Bureau (NCRB) but suffers from bias due to under-reporting and poor registration techniques (The Hindu). Moreover, I use the resources of the National Census on India, for population data, and of the Planning Commission, for data on GDP. All the data are collected at district level in order to exploit the regional variation between districts.

Indian data are used to calibrate the model and exactly identify its endogenous parameters to obtain a distribution of values at district level for education quality (ψ), education cost (ρ) and crime deterrence (δ). After the calibration, I simulate the model under different settings, or counterfactual experiments, to study the evolution of education, savings, human capital

and crime when the parameters of interest vary along their distribution. A simulation where in all districts ρ is set at its 5th percentile value is used to study the quantitative implications of an exogenous decrease in the cost of education access. Setting ψ at its 95th percentile value or δ at its 5th percentile value shows the consequences of a positive exogenous shift in the quality of the education system or in crime deterrence technology. The goal of these numerical exercises is to evaluate the quantitative implication of my theory, that is to provide an idea of the size of the effects on crime deterrence of the mechanisms described in the model. As the changes in the parameters are exogenous and come at no costs, this is not a welfare evaluation.

The research question of this paper stems from a wide scientific literature debating on the potential of education as a tool of crime deterrence. It is largely recognized that education increases the opportunity cost of crime. For example, Naci Mocan, Billups and Overland (2005) present a dynamic model of crime in which every individual possesses two types of human capital, legal and illegal, arguing that changing the relative returns of these capitals in favor of the legal one, is essential for the individual to drop criminal activities. A way to change these relative returns is education (Lochner, 2004, 2007). There are three channels, market and non-market related, through which education can reduce crime incentives: the income effect, more education means more income and so higher return from the legal activity; the time availability effect, more education means longer time spent in school and less time available to commit crime; and the change in patience and risk aversion as educated agents value the future and avoid risky criminal present gains (Witte and Tauchen, 1994; Usher, 1997; Fajnzylber et al., 2002; Lochner and Moretti, 2003; Buonanno and Leonida, 2006, 2009; Machin, Marie and Vuijic', 2011).

The negative association between education and crime can also be observed in the data. In tab.2.1 I regress the logarithm of the number of crime per person in 2011 (namely total crime in model 1, crime against property in model 2 and homicides in model 3) over some education indicators, using districts as units of analysis. I use the gross enrollment rate in primary and upper primary schools (Total GER) as a proxy of the accessibility of education, and the pupil teacher ratio in primary and upper primary schools (Pupil Teacher Ratio) as a proxy of the quality of education. In this way I can account for two important dimensions of education. The gross domestic product per capita and a dummy for the state, are used as controls. The coefficients show that there is a negative association between education indicators and the number of reported crimes. That is within a given state districts investing more in education outperform the others in terms of crime reduction: crime drops when the number of children in school increases or the number of children per teacher decreases, hence

assuring higher education quality.

Table 2.1: Linear Regression of Crime Rate over Education in 2011

	(1)	(2)	(3)
VARIABLES	Model 1	Model 2	Model 3
Total GER	-0.00220	-0.00499**	-0.00338*
	(0.00143)	(0.00225)	(0.00173)
Pupil Teacher Ratio	0.00518***	0.0110***	0.00339**
	(0.00128)	(0.00202)	(0.00155)
GDP per Capira in rupee	2.95e-06***	4.91e-06***	1.27e-06**
	(4.22e-07)	(6.65e-07)	(5.12e-07)
Observations	504	504	502
R-squared	0.638	0.437	0.257

Standard errors in parentheses

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

Dummy for state are omitted in the table

However, education can also encourages crime, if in the society there is inequality in the distribution of income, as it increases the amount of resources to be stolen (Ehrlich, 1973, 1975). According to Ehrlich (1973, 1975) and Becker (1968,1993) in fact, criminals respond mainly to economic incentives: when income inequality is high, potential criminals have little to lose and expect high gains from crime. In this scenario, education expansion, which increases income per capita (Glomm and Ravikumar, 1992), must be coupled with education equality to avoid an increase in the polarization of the society, which leads to frustrated aspirations and possibly to more crime and violence (Ehrlich, 1973, 1975; Ray, 2003; Appadurai, 2004).

Until now I have assumed that education changes the incentives of potential criminals, but not their preferences about risk, time or social issues. This is a common practice in the economic literature, where individuals face decision problems with the objective of maximizing their utility, everything else held constant. In this paper, while acknowledging the simplification brought by this assumption, I will also consider individual preferences as constant in order to better reach the aim of any theoretical work, which is to maximize the explanatory

power and relevance, in terms of policy evaluation, of an economic model (Bowles, 1998; Becker et al., 2012; Becker, 2013). Considering a more complex point of view however, education can also contribute to reduce crime by changing the values, perceptions, beliefs and the morality or the considerations of individuals. As Becker explained (1993), a society in which people, just by following their ethical and moral code do not choose crime, even if it is profitable or with no risk, makes police or jail unnecessary. Education could therefore be used to convey new values and beliefs such that preferences towards crime change. Even if this is not addressed here, it might be the focus for future extensions of the model.

To conclude, in the literature there is no theoretical model addressing the contrast between the cost opportunity of human capital and the " cake size effect ", both originated and nourished by the process of human capital accumulation. In most cases when education is an endogenous equilibrium result, crime is not included in the modelization strategy (Glomm and Ravikumar, 1992; Lesthaeghe et al., 2005). However where crime level is endogenous, the focus is more on the effects of deterrence policies or incapacitation measures, like imprisonment or sanctions, on illegal activities (Becker, 1965; Naci Mocan, Billups and Overland, 2005), without directly addressing education as a trigger of crime reduction. Only Lochner (2004) develops a model where individuals choose how much time to dedicate to education versus criminal activities, and where both these variables are endogenous. However, his model considers individual decisions as based only on individual characteristics (skill levels, learning ability, criminal ability) and not on aggregate wealth, thus leaving no space for the cake size effect. Likewise Fella and Gallipoli (2008) present an overlapping generation model where education and crime are endogenous. However, their focus is not on the mechanisms behind the individual decisions on education and crime, but rather on evaluating how much a policy intervention on high school completion can be cost-effective and cost-viable with respect to a policy intervention on the punishment system.

In order to contribute to the debate, this paper presents a general equilibrium model where both education and crime are endogenous choices, taking into account both individual and aggregate dimensions, and generating some effects with significant quantitative implication for crime levels in the society. As the results of this model point out, when the cake size effect is taken into account, positive shifts in education accessibility or productivity, which boost the accumulation of human capital and exploit the cost opportunity mechanism, are not enough to reduce the level of criminal activities in the long run.

2.2 The Model

Consider an overlapping generations model with three periods. Agents are children in period 1, adults in period 2 and old in period 3. Their utility depends on their retirement income (d_{t+1}), or savings, and on the human capital developed by their children (h_{t+1}). Parameter γ represents the taste for the education level of the progeny.

$$U_{d,h} = d_{t+1} + \gamma \log h_{t+1} \tag{2.1}$$

In both periods 1 and 3 agents are not productive. When they are young they go to school, when they are old they enjoy their savings but have no other income. In period 2, as adults, they decide whether to work legally or illegally (occupation choice) and how much of their income to invest in the education of their children (education investment choice). On the one side, the occupational choice affects the elderly: if adults choose to be criminal, they will steal from the savings of the old people and so reduce their retirement income. On the other side, the investment in education choice affects the children: the more adults invest in education the less they save, the higher will be the human capital of their children as adults. Therefore, the more is invested by parents in the education of their kids, the higher will be the opportunity cost for the children to shift to the illegal sector once they grow up. However, as the human capital per capita increases, also the total productivity of the society and its saving capacity will increase.

Honest adults receive a legal income correspondent to their human capital plus a lump sum transfer (ω) and net of the cost of education, which is given by the investment in schooling (e_t) multiplied by the education accessibility parameter (ρ) (eq.2.2). Criminals enjoy the criminal gain (g_t), the size of which is determined by what is stolen from the elderly and the transfer (ω). Moreover, as they have decided to be criminal, they loose all their human capital and they pay the good cost of crime η_i, which is heterogeneous and uniformly distributed over the interval $(0; 1)$. Finally, they also pay the education investment for their children (eq.2.3).

$$y^{il} = h_t^i - \rho\epsilon_t^i + \omega \tag{2.2}$$

$$y^{icr} = g_t - \rho\epsilon_t^i - \eta_t^i + \omega \tag{2.3}$$

In the following I will consider the occupational and the education choices of the households in a private education framework, that is the burden of the education cost is entirely on the

parents. At equilibrium I am interested in the achieved level of human capital, as a proxy of the incentive to stay in the legal sector, and in the share of criminals in the society.

2.2.1 The Education Choice

At $t = 0$ each household consists of one agent for each generations (young, adult, old). The old agent is endowed with a stock of savings (s^i_{-1}), while the adult agent can use his stock of human capital ($h^i_0 > 0$). The adult agent chooses his occupation, then allocates his income between the education investment for his child and the savings for his retirement. The maximization problem of the adult agent is the following:

$$\max_{\epsilon^i_t} d^i_{t+1} + \gamma \log h^i_{t+1} \tag{2.4}$$

s.t.

$$h^i_{t+1} = \psi(\epsilon^i_t)^\theta h^\nu_t \tag{2.5}$$

$$d^i_{t+1} = Rs^i_t(1 - \delta I_{t+1}) \tag{2.6}$$

$$s^i_t = \begin{cases} h^i_t - \rho\epsilon^i_t + \omega, & \text{if honest} \\ g_t - \rho\epsilon^i_t - \eta^i_t + \omega, & \text{if criminal} \end{cases} \tag{2.7}$$

Households choose the investment in education ϵ^i_t to maximize their utility. They are better off if their retirement income is higher (eq.2.6) and if the human capital of their children is higher (eq.2.5). The fact that their utility is directly defined over the level of human capital of their siblings can be explained by several factors. Firstly, agents might consider higher human capital as good for signaling in the job market, or they simply might care for the relative position of their descendants. Secondly, it can be explainable as a sign of altruism of older generations towards the younger ones. Another reason is that adults might simply consider the young generation human capital as a potential support for old age, thus making the investment in education a substitute of savings. We will see later that this is a key mechanism for the equilibrium results in terms of education and crime. Whatever the household is able to save is used in the retirement period and is exposed to crime (I_{t+1}) perpetuated by the new criminals (eq.2.6). The low of accumulation of human capital (eq.2.5) depends on the productivity of the education system (ψ), on the investment in education (ϵ) and on the average human capital accumulated by the generation of the parents(h_t). θ is the rate of return of the parental investment in education, while ν is the rate of return of parental human capital.

The F.O.C. of this problem is

$$\epsilon_t^* = \frac{\gamma\theta}{\rho R(1-\delta I_{t+1})} \tag{2.8}$$

The optimal education investment is independent on the level of human capital of the parents. If I replace the optimal value of ϵ_t in the production function of human capital (eq.2.5) I obtain the equation for the optimal level of young adults human capital:

$$h_{t+1}^* = \psi(\frac{\gamma\theta}{\rho R(1-\delta I_{t+1})})^\theta h_t^\nu \equiv h(I_{t+1}, \psi, \delta, \rho) \tag{2.9}$$

Eq.2.9 shows that the human capital of the children depends positively on the expected crime level for the next period. This is the substitution mechanism, the first of the three mechanisms that can be observed in the relationship between education and crime. As the utility of adult agents depend both on their retirement income and the human capital level of their children, if parents expect a high crime rate in their retirement period they decide to allocate more resources to the education of the children and less to savings, to be less exposed to crime.

2.2.2 The Occupation Choice

The occupation choice of the adults depends on their human capital and on the cost for the criminal activity. I look at the choice from the point of view of the agent adult in t. Given the legal and illegal income:

$$y^{il} = h_t - \rho\epsilon_t + \omega \tag{2.2}$$

$$y^{icr} = g_t - \rho\epsilon_t - \eta_t^i + \omega \tag{2.3}$$

from $y^{icr} = y^{il}$ I can find the threshold $\tilde{\eta}$ such that an individual is indifferent between the illegal and the legal sector.

Proposition 2.1: *at each period t there exists $\tilde{\eta}_t = I_t = g_t - h_t$ such that for all individuals with $\eta^i < \tilde{\eta}$ then $y^{icr} > y^{il}$ and they will prefer to stay in the criminal rather than in the legal sector; while for all individuals with $\eta^i > \tilde{\eta}$ then $y^{icr} < y^{il}$ and they will prefer to stay in the legal rather than in the criminal sector .* [1]

[1] From (2.2) = (2.3) I obtain an expression for $\tilde{\eta}$. However $\tilde{\eta}$ is also the share of criminals in the population at time t ($\tilde{\eta}_t = I_t$) due to the good cost of crime is following a uniform distribution.

The share of criminals in a given period depends therefore on the human capital accumulated when young and on the criminal gain. Higher is the human capital an individual is endowed with, higher is the opportunity cost to switch to the illegal sector; while higher is the criminal gain, higher is the incentive the individual has to join the illegal sector. The criminal gain of a given period is the sum of what is stolen from the savings of the old agents, both honest or criminals, divided by all the criminals in the society in that period (eq.2.10). Eq.2.11 shows the aggregate savings for a given period t.

$$g_t = \frac{\delta I_t R S_{t-1}}{I_t} = \delta R S_{t-1} \tag{2.10}$$

$$S_t = I_t(g_t - \rho\epsilon_t + \omega) - \int_0^{\tilde{\eta}} \eta_i G(\eta_i) d\eta_i + (1 - I_t)(h_t - \rho\epsilon_t + \omega) \tag{2.11}$$

Criminals are able to steal more if they are more productive. The productivity of crime, which can also be called the crime deterrence technology, is δ. Higher (lower) δ means that crime is more (less) productive, and deterrence measures are less (more) effective.

2.2.3 Temporary Equilibrium

The core results of the education choice and occupation choice are the level of human capital of young adults (h_{t+1}) and the crime share (I_t/I_{t+1}) such that:

- the education, or portfolio choice, is represented by (eq.2.8 and 2.9)

$$\epsilon_t^* = \frac{\gamma\theta}{\rho R(1 - \delta I_{t+1})} \tag{2.8}$$

$$h_{t+1}^* = \psi(\frac{\gamma\theta}{\rho R(1 - \delta I_{t+1})})^\theta h_t^\nu \tag{2.9}$$

- the occupation choice for adults in t is given by (eq.2.12), while (eq.2.13) expresses the occupation choice for adults in $t + 1$.

$$I_t = \delta R S_{t-1} - h_t \tag{2.12}$$

$$I_{t+1} = \delta R S_t - h_{t+1} \tag{2.13}$$

Looking at eq.2.9 and eq.2.13 it is possible to identify three mechanisms at play in the education-crime dyad.

First the investment in education, and therefore in the level of human capital of the siblings, is positively related to the expected level of crime in the following period. The education investment is a substitute of savings to be less exposed to crime during the retirement age. I will call this the substitution effect.

Second the share of criminals in the society in a given period is negatively related to the level of human capital of the adult generation of that period: higher is the human capital of adults, higher will be the income they lose if they choose to switch to the illegal sector, that is higher will be the opportunity-cost to become criminals. I will call this mechanism the opportunity cost of education.

Finally the number of criminals in the society is positively related to the level of aggregate savings of the old generation. Higher is the level of savings, higher is the amount of resources that can be stolen. This is the cake size effect.

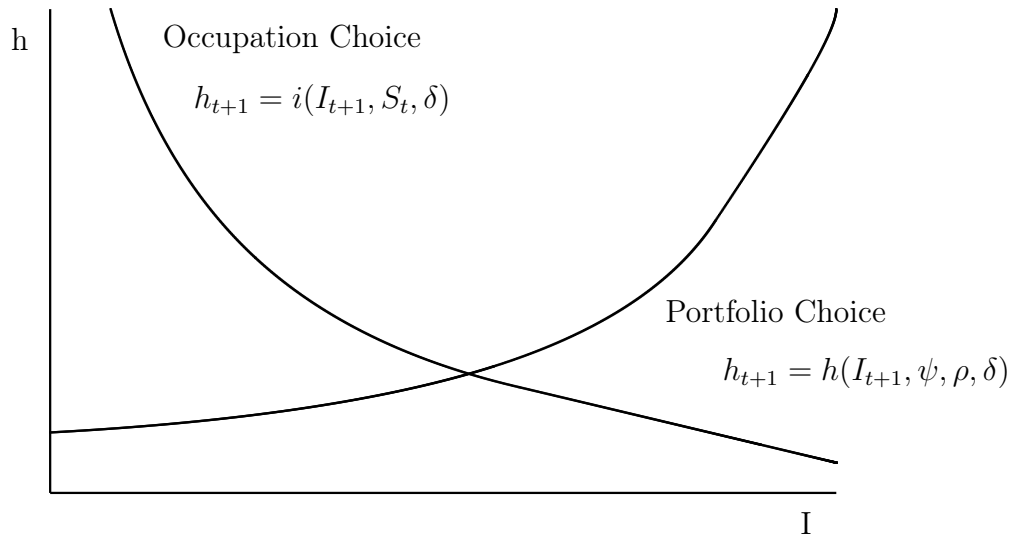
From (eq.2.13) we can express the occupation choice with human capital as a function of crime, savings and the other parameters in order to represent the occupation and the portfolio choice graphically.

$$h_{t+1} = \delta R S_t - I_{t+1} \equiv i(I_{t+1}, S_t, \delta) \tag{2.14}$$

To be clear I_{t+1} is expected at the time of the education choice of the generation of adult in period t, however it materializes in $t+1$, at the time of the occupation choice of the adult generation in t+1, as a result of an aggregation of behavior from the previous period. [2]

[2]The temporary equilibrium should be ideally expressed in terms of h_t and I_t. However in the model the choice of education and saving in t are strictly connected to the level of crime in $t+1$ due to the substitution effect. I_{t+1} is the result of an aggregation of behavior in t. That is the result of the choices in terms of human capital accumulation and crime are visible only in $t+1$. For this reason the temporary equilibrium is expressed in term of $t+1$.

Figure 2.1: Temporary Equilibrium



Proposition 2.2 (Temporary Equilibrium): *An equilibrium with private education funding is a vector of young adults human capital and criminal share* $(h_{t+1}, I_{t+1})_{t\geq 0}$ *satisfying eq.2.9 and eq.2.13. It exist and it is unique. It is such that:*

1. *the level of human capital of young adults (eq.2.9) is increasing in* ψ *and* δ *and decreasing in* ρ*;*
2. *the level of crime in the society (eq.2.13) is increasing in* δ*.*

See appendix A and appendix B.

The portfolio choice curve is upward sloping and depends on the parameters - δ, ψ and ρ- in addition to crime (I_{t+1}). A shift in the parameters generates a shift of the portfolio choice curve in the plan, while a change in the expected level of crime originates a change along the curve. When crime deterrence is less effective (high δ) or high crime is expected in the next period (high I_{t+1}), the substitution mechanism comes into play and agents invest more in education to avoid to be exposed to crime. Human capital increases also if the education system is more productive (high ψ), while decreases if the education investment becomes more costly (high ρ).

The occupation choice curve is downward sloping and depends on the parameter δ and on S_t and h_{t+1}. A shift in δ or in S_t generates a shift of the curve in the plan, while a change in h_{t+1} originates a change along the curve. When crime is more effective (high δ) or there

is more to be stolen (high S_t), the criminal gain increases as well as the incentive to join the illegal sector. When young adults have high human capital (high h_{t+1}) they face a costly trade off between the legal and the illegal sector and are less motivated to became criminals.

The combined effect of the parameters on the two curves however, is not so straightforward to predict. For example, if a shift in δ makes human capital increasing much more than crime incentive (the portfolio curve shifts more than the occupation choice curve), crime drops at the new equilibrium. However, if crime incentives react more than the education investment to a change in δ, the opposite will happen. The final result in terms of human capital and crime depends therefore on the steepness of the two curves and on the relative response to a change in the parameters. It might be that, due to the substitution mechanism, a decrease in crime deterrence technology ($\delta \uparrow$) generates a fall of crime rates instead of an increase. Moreover, a drop in ρ or an increase in ψ, may also have a controversial result. First, they generate an increase in human capital, which increases the cost opportunity of young adults in the short run. However, in the long run, also aggregate savings increase boosting the cake size effect and crime incentives. Positive changes in education accessibility or productivity may therefore work in the short run but not in the long run.

For this reason in section 2.3.2 I will calibrate the model to exactly identify this set of parameters at district level in India. Once I have the calibrated parameters, I will conduct counterfactual experiments to study the effect of exogenous technological changes in education practices and crime deterrence on the equilibrium values of human capital and crime. The aim of this quantitative check is also to see if crime reacts to technical changes in a range of reasonable values of the parameters.

2.3 Quantitative Analysis

2.3.1 Data

To calibrate the model and implement a quantitative analysis, I need the observed values of human capital, education investment, savings and crime against property. These data are gathered from different sources. I used the INDIAN CENSUS for population aged 6-14 and total population, to calculate the enrollment rates and per capita productivity measures; the National Crime Records Bureau (NCRB) for crime data at district level; the DISE for enrollment data in primary and upper primary schools; and the Planning Commission databases for total GDP at district level. Census data on population were available for two

years, 2001 and 2011, therefore 2001 observed data is " Period t", while 2011 observed data is " period $t+1$". Unfortunately, I had data for GDP at district level for both the census years only for 280 districts out of 686, so I had to reduce the sample available. Among many, the districts included in the analysis are part of the following states: Andhra Pradesh, Assam, Bihar, Karnataka, Kerala, Maharashtra, Orissa, Rajasthan, Uttar Pradesh and West Bengal.

Data on education are taken from the DISE database. To obtain the gross enrollment rate (GER) for year 2001 and 2011 I sum the total enrollment, that is the number of children enrolled in school, from grade I to grade VIII and I divide this sum by the total number of children between 6 and 14 years as registered in the Census of India in 2001 and 2011. Usually children aged 6 to 10, that is in grade I to V, go to the primary school, while children aged 11 to 14, that is in grade VI to VIII , go to the upper primary school. Summary statistics for the GER across districts in 2001 and 2011 are shown in tab.2.2 and in figures 2.2 and 2.3. Values which are higher than 100% are possible given the fact that some children may repeat some years of school and therefore find themselves still in primary or upper primary classes even if they are older than 14. We can observe that the average enrollment increases over the time, however there is a certain percentage of children still out of school. Finally, the GER varies at regional level and over time, showing that the RTE act is being implemented and that I can use this variation for the quantitative analysis which follows.

Table 2.2: Summary Statistics for Gross Enrollment Rate in 2001 and 2011

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	N	mean	sd	min	max	p25	p50	p75
TOTGER	281	63.32	16.99	11.20	180.5	51.69	64.03	73.89
TOTGER_2011	281	81.14	14.88	16.19	165.5	74.04	80.96	88.61

Note: GER is measured in %

Figure 2.2: Gross Enrollment rate in Indian Districts in 2001 (Source: DISE)

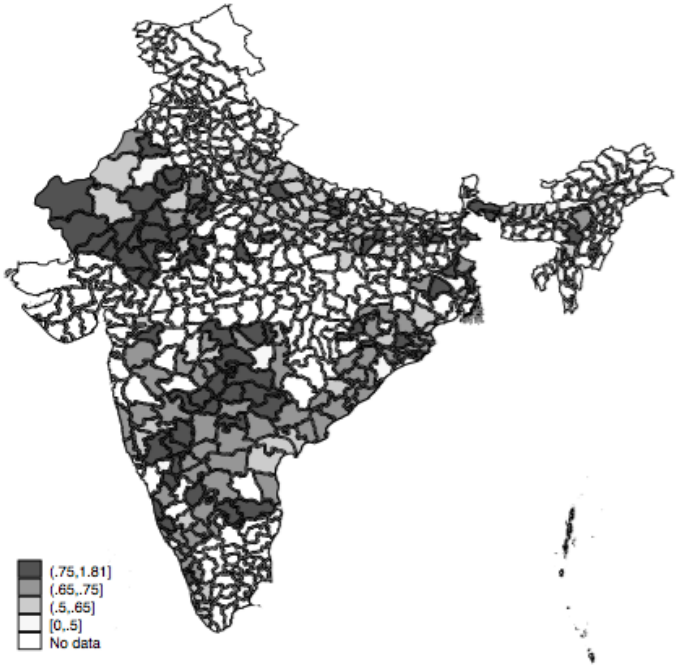
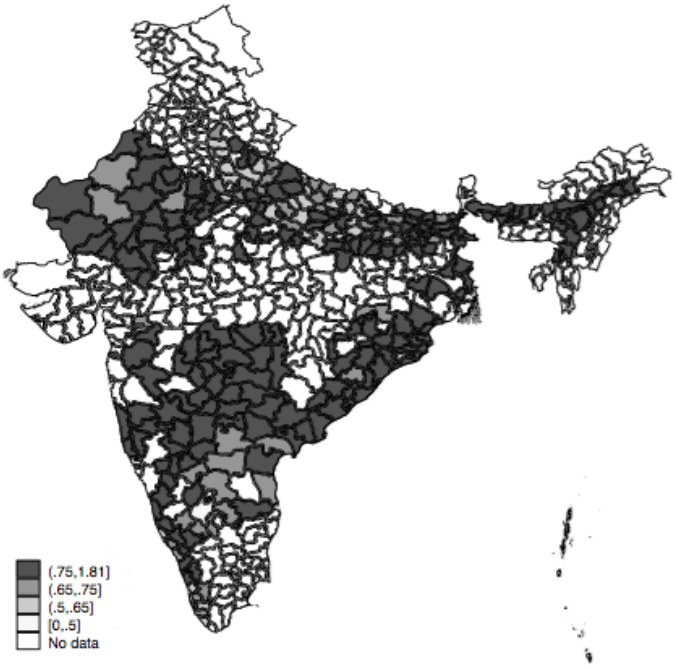


Figure 2.3: Gross Enrollment rate in Indian Districts in 2011 (Source: DISE)



To obtain the initial per capita values of human capital (h_0) and savings (s_{-1}) I start from the GDP per capita at regional level. The GDP per capita of a district is measured dividing the aggregate GDP in rupees by the total population of age between 15 and 60. I additionally

rescale this value of per capita productivity diving it by 1000. I assume that in period t more or less 1/5 of the GDP is the income devoted to savings by the old generation in the previous period. The rest is the adult generation human capital, considering that part of this income is lost by the agents choosing to be criminals. I use this assumption as there are no data available for savings at district level in the years of interest. Human capital and savings are therefore the solution of the following system, where i is the index for the districts.

$$GDP_0^i = h_0^i(1 - I_0) + Rs_{t-1}^i \tag{2.15}$$

$$Rs_{t-1}^i = \frac{1}{5} GDP_0^i \tag{2.16}$$

For crime data I use the only source available in India, the National Crime Record Bureau. In the NCRB database there are many different types of crime for which I have detailed information until the district level. Crime is generally classified as crime against person, crime against property, riots, crime against children and sex crime. My analysis focuses on property crime for two main reasons. Firstly, property crime is usually considered a blue collar crime, meaning that the skills required to perform such criminal activities are not related to an increase in education. Secondly, working with crime data in India is a delicate matter as the number of crime acts registered is exposed to bias due to under-reporting and poor registration techniques (The Hindu). However, property crime and homicides are usually the highest quality data among all the other criminal imputations. The district values of property crimes in percentage, for 2001 and 2011, are shown in tab.2.3 and figures 2.4 and 2.5. To calculate the crime rate firstly I multiplied the number of crime against property reported in each district by two, due to the under-reporting issue, then I multiplied the figure obtained by 10, to take into account that 1 period lasts ten years, from 2001 to 2011 and from 2011 to 2021. Finally I divided the result by the total number of individuals of age between 15 and 60 in that district.

Table 2.3: Summary Statistics for Crime Rate at District Level in 2001 and 2011

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	N	mean	sd	min	max	p25	p50	p75
Cr_2001	281	1.39	1.12	0.189	7.87	0.703	1.04	1.70
Cr_2011	281	1.36	1.43	0.0104	18.5	0.687	1.03	1.61

Note: Crime is measured in %.

Figures are calculated tanking into account under-reporting problems.

Each period under analysis lasts 10 years.

Figure 2.4: Crime Rate in Indian Districts in 2001 (Source: NCRB)

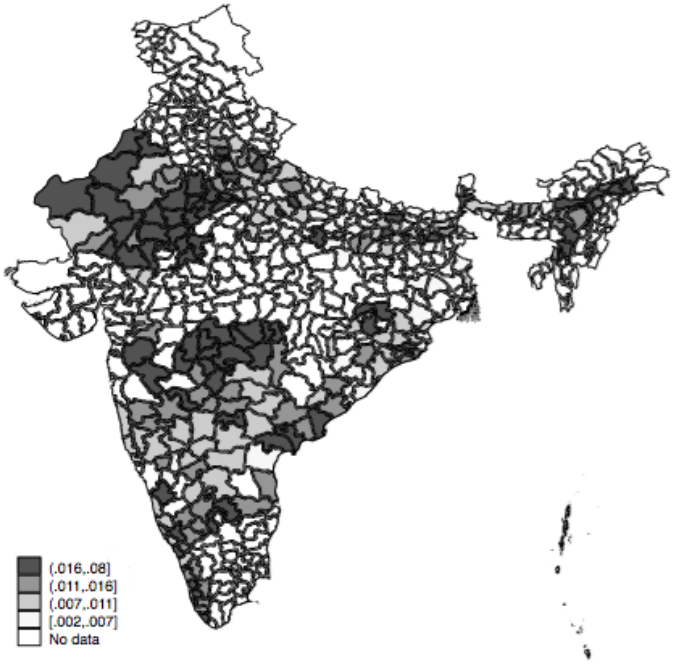
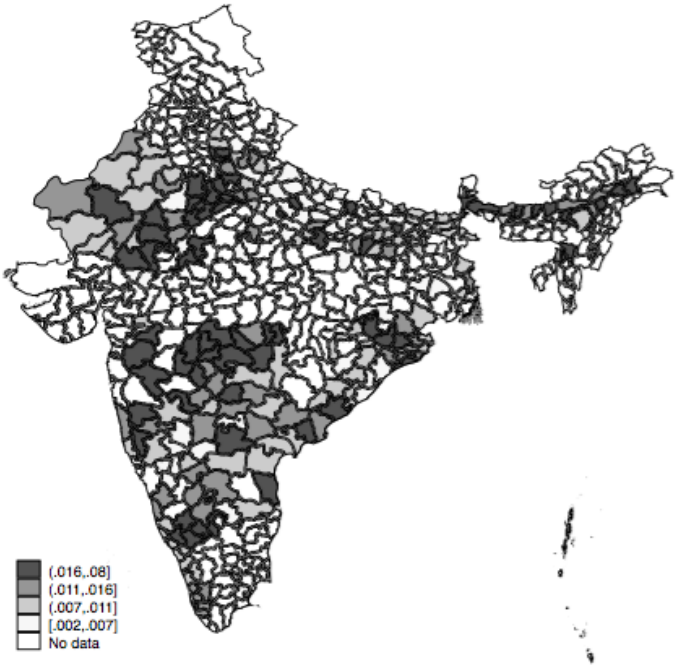


Figure 2.5: Crime Rate in Indian Districts in 2011 (Source: NCRB)



2.3.2 Calibration

To be able to pursue a quantitative analysis, in addition to the observed values for human capital, education, savings and crime against property (tab.2.4), I need to assign a value also to the other parameters. A first set of parameters are fixed a priori. In line with the literature the saving rate (R) is 1.02^{10}. The rate of return of parental investment in education θ is 0.5 and the rate of return of parental human capital is 0.1 (tab.2.4). These parameters are taken from De la Croix and Doepke (2003).

The education taste parameter γ is equal to 60. In principle the value of γ should be chosen such that the share of education spending over GDP of the model matches the one observed in the data. For India the latest data available on the Indian Data Portal on education spending are for 2013 and correspond to 3.8% of GDP. However I did not follow this method as, by construction, in the model there are only two possible kind of GDP expenditure: education or savings. In section 2.3.1 I set savings in period -1 as $\frac{1}{5}$ of the observed GDP. This assumption is needed to have the district level of savings for period 0. The rest of household spending is devoted to education by construction. Setting γ such that the share of education expenditure on GDP matches the observed one does not make sense considering how the model is built and considering the dynamics investigated in this analysis. To have such a small value it would mean that in period 0 almost everything is devoted to savings

and nothing to education. As a matter of fact, the calibration is not too sensitive to the choice of γ. I have tried to calibrate, and then simulate, the model for γ equals to 60, 30 and 90, as robustness checks. The values of ψ and δ do not change, while ρ and ω adjust. However, the share of education spending on GDP remains in the range of 83% to 96%. Given that the calibration and simulation are not so sensitive to γ, I set it at the central value of 60 (Appendix C).

Table 2.4: Summary Statistics of the Exogenous Parameters and Indian Data

Exogenous Parameters	**Description**	**Assigned Value**
R	Interest Rate	1.02^{10}
γ	Taste for educated progeny	60
θ	Rate of return of parental investment in education	0.5
ν	Rate of return of parental human capital	0.1
n	Cost-opportunity of Crime	1
Variables of Interest	**Description**	**Districts Average in Data**
Hc_{2001}	Individual Human Capital in Rupees in 2001	1.88
Hc_{2011}	Individual Human Capital in Rupees in 2011	3.48
S_{1990}	Individual Savings in Rupees in 1990	0.38
S_{2001}	Individual Savings in Rupees in 2001	0.70
Cr_{2001}	Crime rate for population at risk in 2001	0.014
Cr_{2011}	Crime rate for population at risk in 2011	0.014
ϵ_{2001}	Gross Enrollment Rate	0.63

A second set of parameters is exactly identified at district level calibrating the model. To calibrate the endogenous parameters, I use the following system of equations:

$$\epsilon_t = \frac{\gamma\theta}{\rho R(1-\delta I_{t+1})} \tag{2.17}$$

$$I_{t+1} = \delta R S_t - h_{t+1} \tag{2.18}$$

$$S_t = I_t \delta R S_{t-1} + \omega + h_t(1-I_t) - \frac{I_t^2}{2} - \rho\epsilon_t \tag{2.19}$$

$$h_{t+1} = \psi \epsilon_t^{\theta} h_t^{\nu} \tag{2.20}$$

Eq.2.17 represents the optimal level of education investment. Eq.2.18 is the crime share resulting from the choice of young adults. Eq.2.19 represents the aggregate savings, while eq.2.20 describes the law of accumulation of human capital. Therefore δ, ρ, ψ and ω are the four unknown of a system of four equations. Summary statistics for the parameter values obtained with the calibration are given in tab.2.5, while appendix D shows their distribution across districts.

Table 2.5: Calibrated Parameters Statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	N	mean	sd	min	max	p05	p25	p50	p75	p95
ψ_i	281	4.105	2.562	0.0706	12.76	0.11	2.744	3.666	5.419	9.16
ρ_i	281	46.92	34.44	14.47	487.8	29.77	35.16	40.60	50.23	67.71
δ_i	281	4.130	0.173	4.001	5.056	4.02	4.039	4.064	4.114	4.53
ω_i	281	26.16	20.97	22.08	375.7	23.66	24.25	24.65	25.22	27.09

Note: Subscript i stands for districts as units of analysis.

2.4 Simulation and Counterfactual Experiments Results

With the calibrated regional values for the technical parameters I simulate the model to study the equilibrium values of education, savings, human capital and crime under different settings. The first step is to simulate the model with district values for the core technical instruments to have a benchmark simulation at the regional level. In this case each district faces a different set of parameters as a result of different socio-economic conditions. I then compare these results to counterfactual experiments where I introduce exogenous technological changes in productivity, education cost and crime deterrence. These changes show what would happen if all districts were brought at the frontier at once, all adopting the best or worst practices in terms of education or crime deterrence. Setting ψ to its lowest (5th or 25th percentile) or highest (75th or 95th percentile) value in all districts, along the distribution of values of the parameter across districts, is similar to simulating a shift in the productivity of the education system, as if all regions adopt the worst or the best practice for

schooling. In the same way setting ρ or δ to their lowest or highest values, is like simulating a shift in education accessibility costs or in crime deterrence technology across the districts of India. As I explained the education-crime relationship is not straightforward, therefore the outcomes of crime share and human capital, for an exogenous change in any of the core parameters, cannot be determined a priori. The results of the counterfactual experiments are shown in the following figures. Proceeding one parameter at time, the figures show the evolution of education, human capital, savings and crime from period one to five, for the benchmark simulation and for the four counterfactual experiments where the best or the worst practices are adopted in all the districts.

Figures 2.6 to 2.9 show that when education access increases (ρ at its 5th or 25th percentile) education investment, human capital accumulation and savings react positively and increase with respect to the benchmark. However crime, at least in the short run, is neither decreasing nor increasing. On the contrary in the long run crime reacts to a change in the cost of education. As there is education expansion and more human capital accumulation and savings, there are more aggregate resources. Set in motion by the new wealth, the cake size effect wins over the cost opportunity one and crime rises in the long run.

When school costs increase (ρ is at its 75th or 95th percentile) the opposite occurs. Education is more expensive therefore agents reduce the investment in education. There is less human capital growth over time and young adults have a lower opportunity cost of human capital. As a result in the short run crime increases with respect to the benchmark. However in the long run crime stabilizes at lower levels. As there is less human capital growth and aggregate savings are too low, there are less potential criminals attracted by the illegal sector.

Figure 2.6: Education Investment Evolution (change in ρ)

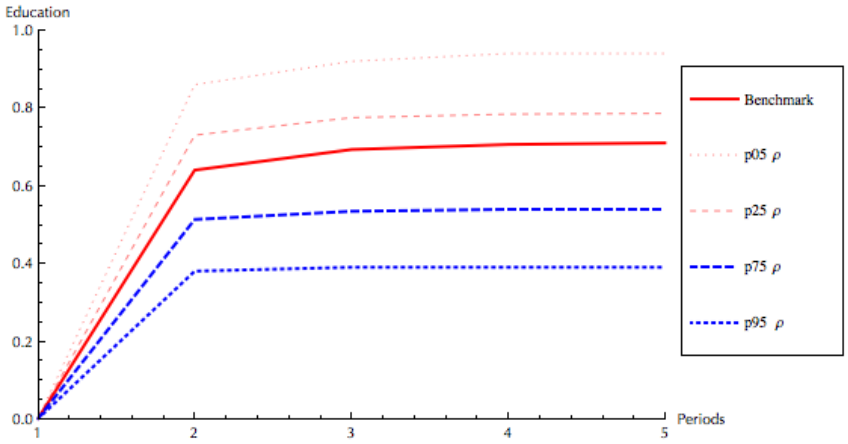


Figure 2.7: Human Capital Evolution (change in ρ)

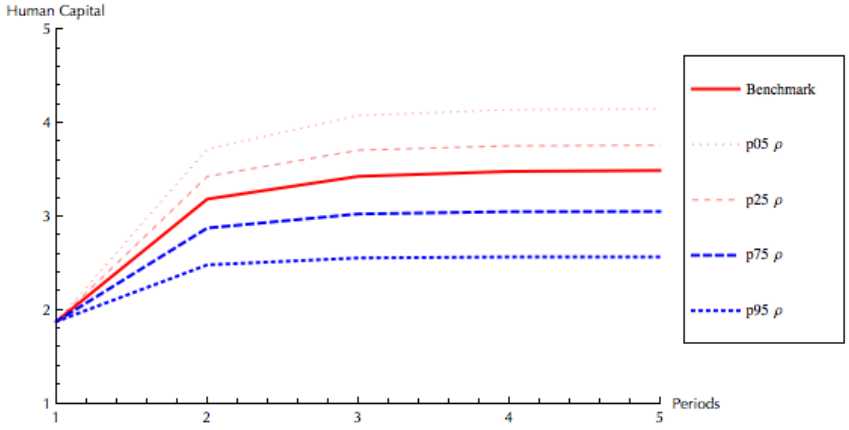


Figure 2.8: Savings Evolution (change in ρ)

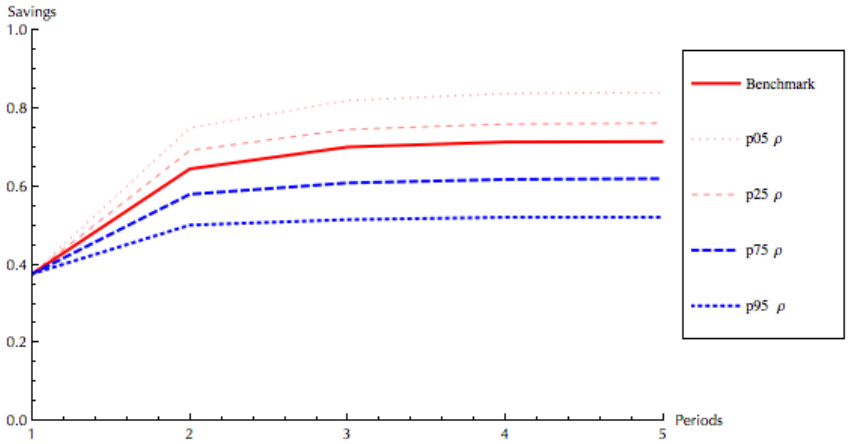
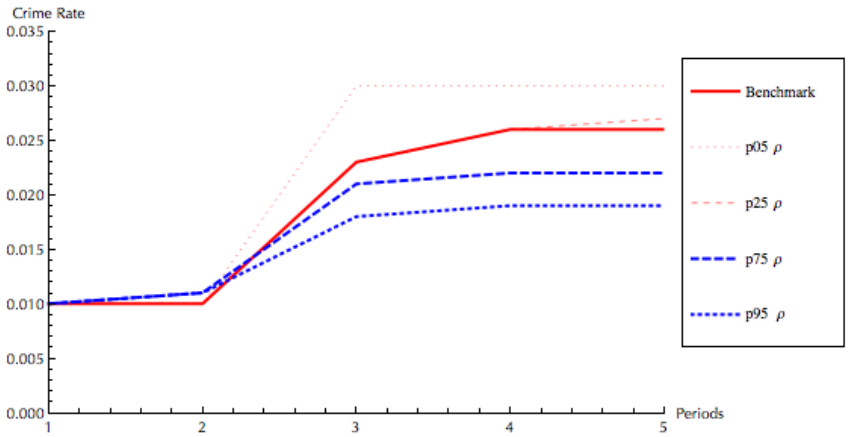


Figure 2.9: Property Crime Rate Evolution (change in ρ)



Figures 2.10 to 2.13 show that the core variables are consistently less sensitive to a change in δ, even if we can observe some variation in the short and in the long run. The most evident change concerns the level of savings when δ is set at its 95th percentile. When δ increases crime is more frightening and criminals are able to steal an higher amount of legal resources. Individuals have less disposable income and save less, not even increasing education as a substitute. This result in crime reduction, especially in the long run.

Figure 2.10: Education Investment Evolution (change in δ)

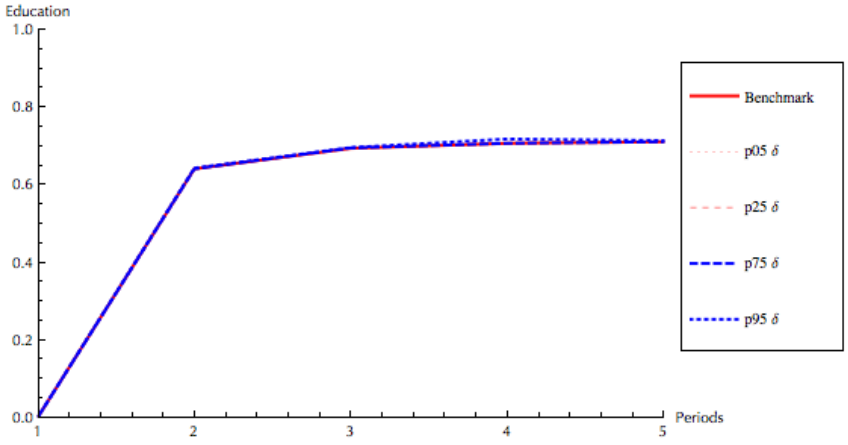


Figure 2.11: Human Capital Evolution (change in δ)

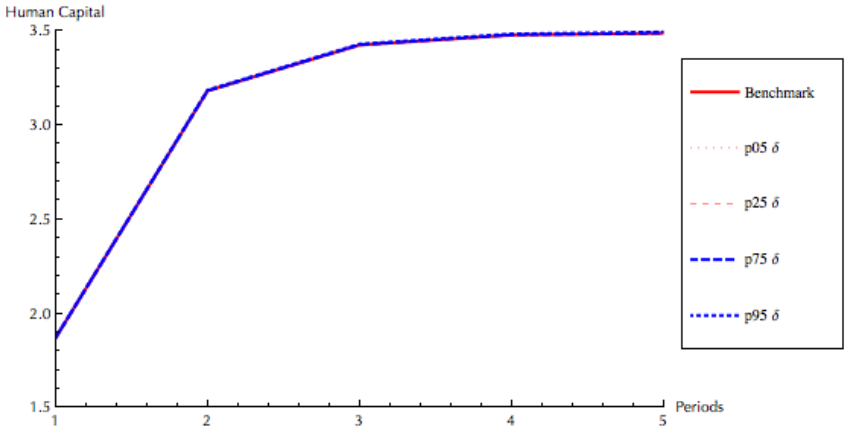


Figure 2.12: Savings Evolution (change in δ)

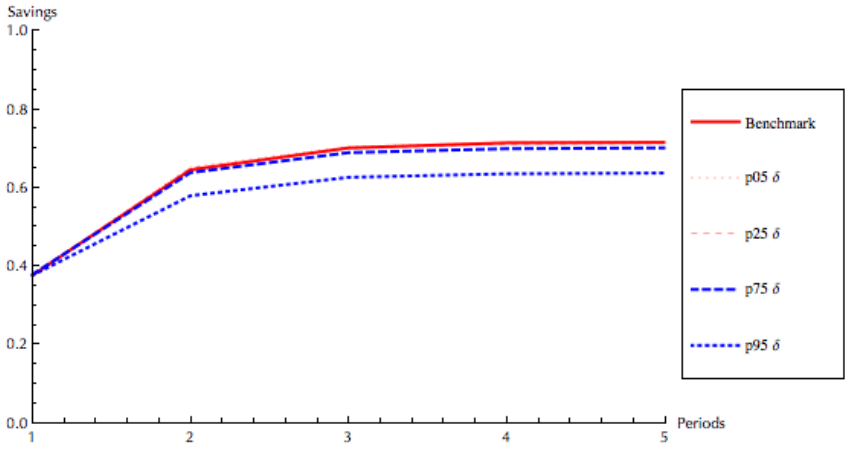
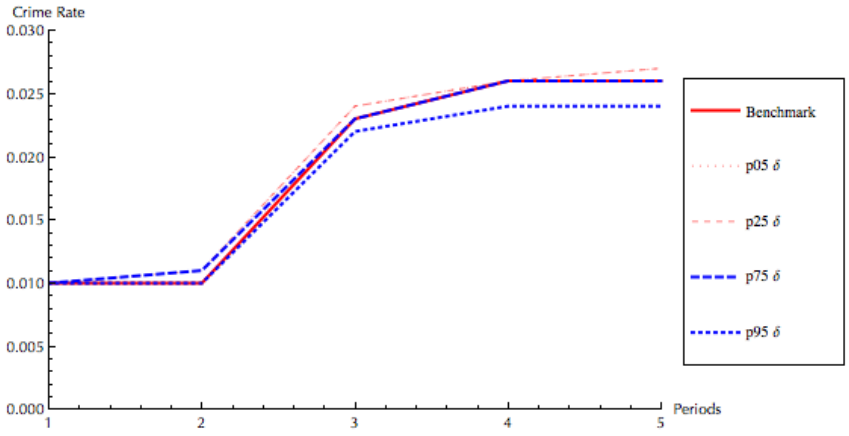


Figure 2.13: Property Crime Rate Evolution (change in δ)



Finally figures 2.14 to 2.17 show that if the productivity of the education system drops (low ψ), crime increases in the short run and then decreases. A negative variation of ψ in fact generates a drop in the level of human capital of the young adults, determining a low cost opportunity of human capital and a high incentive to be criminal. However, later on there is almost nothing to steal and crime drops. On the contrary when ψ increases, human capital also increases for any level of the education investment or savings. Therefore, young adults have high opportunity cost of human capital and low incentives to be criminals. As a result crime drops in the short run. In the long run, due to the high level of productivity, there is high human capital growth which translates also in high aggregate savings and the cake size effect wins.

Figure 2.14: Education Investment Evolution (change in ψ)

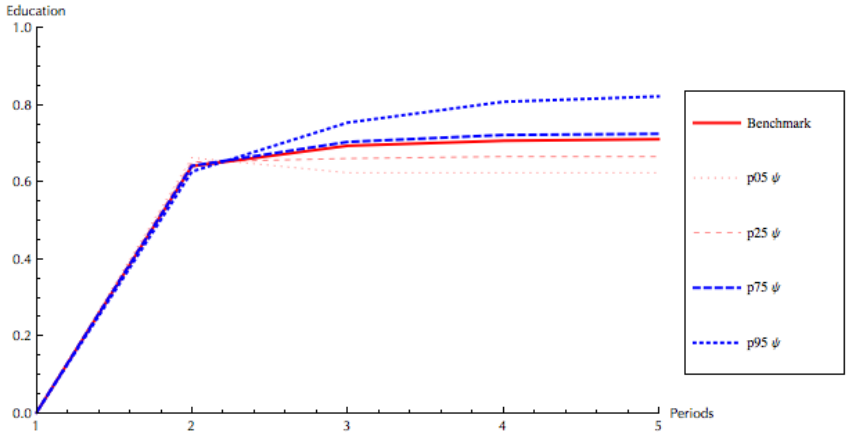


Figure 2.15: Human Capital Evolution (change in ψ)

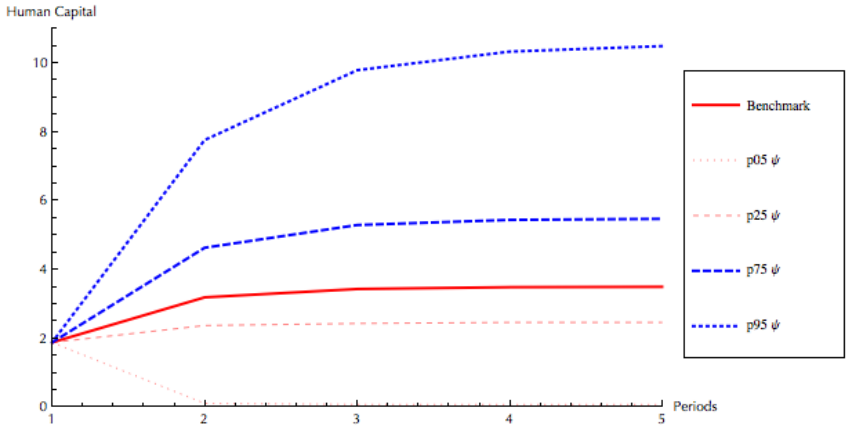


Figure 2.16: Savings Evolution (change in ψ)

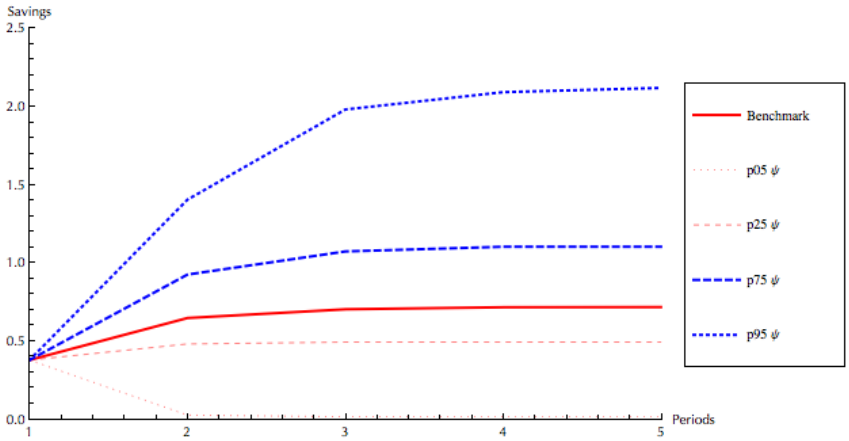
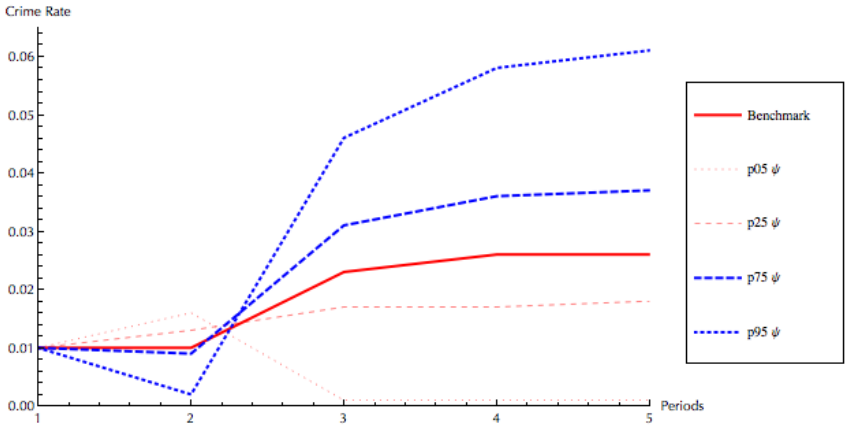


Figure 2.17: Property Crime Rate Evolution (change in ψ)



It is also important to compare results from the quantitative exercises related to different parameters. Fig.2.18 and fig.2.19 show equilibrium levels of crime in the short and in the long run. In fig.2.18 I consider variations in the parameters that boost crime levels, while fig.2.19 represents variations that have a negative effect on crime. As already mentioned, an increase or a decrease in crime deterrence technology has no effect on crime in both the short and the long term, as well as a reduction in the cost of education does not generate benefits in the short run. The results appear more sensitive to a change in ψ: a deterioration of education quality (low ψ) increases crime rates in the short run. However, as this intervention reduces the growth of human capital and savings, there will be a few resources to be stolen in the future such that crime rates are bound to drop in the longer term. As for the short-term, low levels of crime are achieved thanks to an improvement in the quality of the school system (high ψ). Ultimately, wider access to education is not sufficient and efforts need to be made in the quality of the school system to make education a valuable tool against crime.

Figure 2.18: Crime Boosting Exogenous Changes

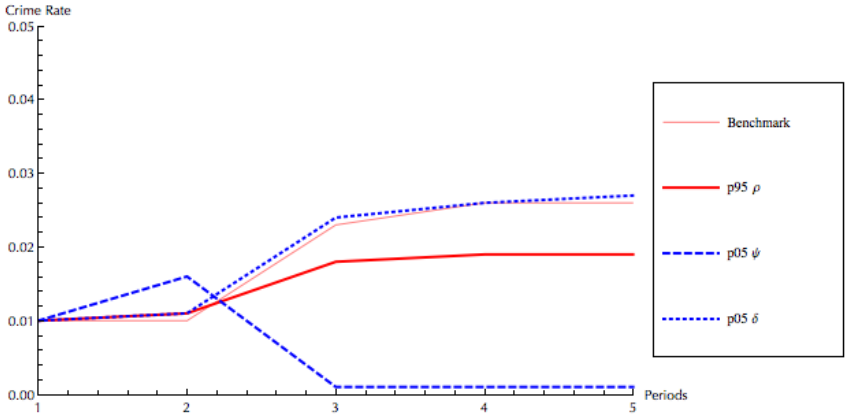
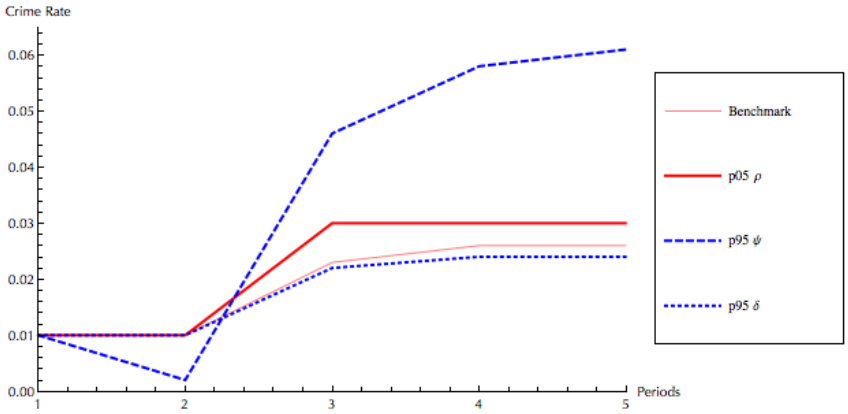


Figure 2.19: Crime Deterrence Exogenous Changes



Indian policy-makers seem to acknowledge the importance of investing in quality of education: The Right for Education Act, for instance, does not only grant free access to education for all children from 6 to 14 but it also establishes strict rules on the quality of the school performance (minimum number of hours at school and minimum number of days at school during one academic year,...), on the quality of the teaching (maximum pupil/teacher ratio or minimum number of teachers in a school,...) and on the quality of the infrastructures (increased number of elementary schools, maximum pupil/classroom ratio, increased percentage of school with water, separate toilet facility, ramp,...).

However, despite the positive outcomes in the short run, fig.2.19 shows that in the long run crime increases in every scenario, meaning that the cake size effect wins. The trigger of this result might be an unequal distribution of income: while getting more people educated increases the returns of legal activity, on the other side this might exacerbate inequality and boost crime incentives, which generate the cake size effect (Ehrlich, 1973, 1975). The persistence in income inequality draws therefore the attention on another aspect related to education provision: the equality of opportunity of the education system, that is the possibility for students coming from a different background to equally succeed at school and in life (Hindricks and Godin, 2016). A society with equal education opportunity will also be a society with a more equal distribution of income (Ehrlich, 1975; Hindricks and Godin, 2016), where individuals are less encouraged to choose the illegal sector when aggregate resources are growing. It is therefore vital to equalize education opportunities and training, to maximize the potentials of good education practices in relation to crime deterrence.

Finally, it is also true that the effect of exogenous changes in the core parameters can vary in different states. It is possible to cluster Indian states on several factors such as

religion, climate, literacy rate, unemployment rate, poverty rate, violence against schedule cast and tribe, or share of schedule cast and tribe. After having tried to classify results according to these different dimensions, I have chosen to show only the results related to poverty rate, that is to the percentage of population living in poverty. In figure 2.20 and 2.21 I have plotted the median crime rate among all districts belonging to states with low($< 15\%$), average (15% to 30%) and high ($> 30\%$) poverty rate. As fig.2.20 shows, in the presence of a high poverty rate an increase in education access is an effective policy to reduce crime in the short and in the long run, with respect to the benchmark. While when the policy is implemented in rich states and districts, there is no effect in the short run and the income effect wins in the long run. As fig.2.21 shows, a policy intervention to increase the quality of education is much more effective against crime in states and districts where there is high poverty. In such cases, coupling education access with education quality policy interventions seems to bring promising positive results.

Figure 2.20: Crime Rate and Poverty Rate (low ρ)

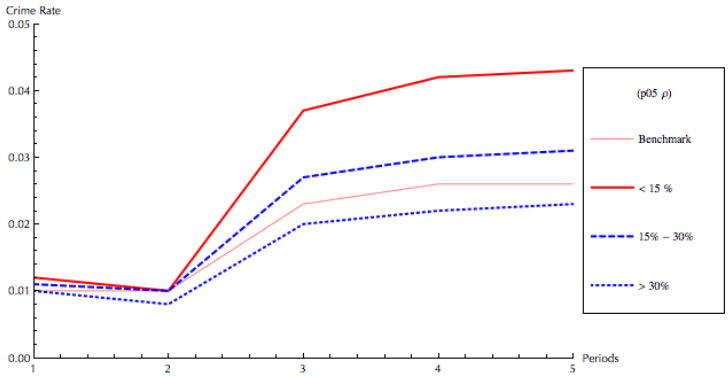
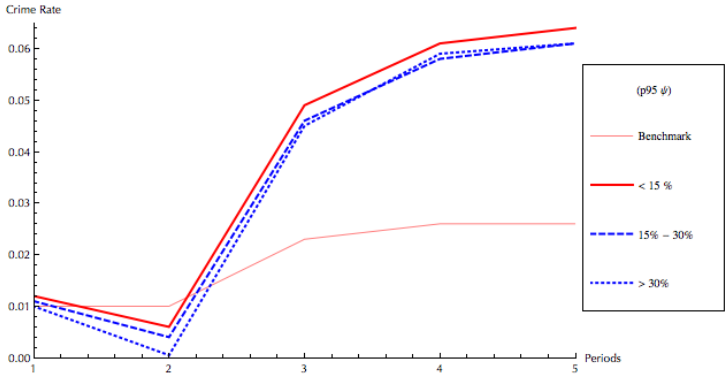


Figure 2.21: Crime Rate and Poverty Rate (high ψ)



2.5 Conclusions

The primary goal of this paper is to determine the conditions under which crime drops with reference to education practices (education cost - ρ - and education quality - ψ -) and crime deterrence efforts (crime deterrence technology - δ -), using a theoretical model where crime and education are endogenous equilibrium results.

To assess the quantitative implications of my theory I implement a numerical analysis with counterfactual experiments, using Indian data on education, crime, production and population, thanks to the rich amount of data available for education in this country. The main results show that to best deter crime in the short run, it is fundamental to tackle both the expansion of education (lowering the cost of access to schooling) and the quality of the school system (more and better schools or better teaching). In the long run however, general equilibrium effects make these policies ineffective. When education boosts human capital growth and productivity, the amount of resources exposed to crime increases as well as the incentives for the agents to choose an occupation in the illegal sector, such that the cake size effect wins. The trigger of this result might be an unequal distribution of human capital and income. The persistence in income inequality draws attention to another aspect related to education provision: the equality of opportunity of the education system. A society with equal education opportunity will also be a society with a more equal distribution of income, where individuals are less encouraged to choose the illegal sector when aggregate resources are growing.

In conclusion, several factors and complex dynamics operate behind the education-crime dyad and to investigate the implications of crime deterrence practices is not an easy task, requiring us to take into account this underlying complexity. Ensuring wider access to education is the first step to make school a tool of crime reduction. Investing in the quality of education is the following one. Increasing also the equality of opportunity of education should also be considered to make schooling effective also in the long run.

2.6 Appendix

Appendix A

Proof of proposition 2.2:

Given S_{t-1}, h_t and I_t, the temporary equilibrium of period t exists and is unique. This equilibrium can be expressed as a function of h_{t+1} and I_{t+1}.

First I want to see how h_{t+1} varies with I_{t+1}. I know h_{t+1} is a function of I_{t+1}, with h_t given. Therefore I can express eq.2.9 in a implicit way as $g(I_{t+1}, h_{t+1}) = 0$.

$$g(I_{t+1}, h_{t+1}) = h_{t+1} - \psi(\frac{\gamma\theta}{\rho R(1-\delta I_{t+1})})^{\theta} h_t^{\nu} = 0 \tag{2.21}$$

The implicit function theorem claims that I can determine how human capital varies with crime using the partial derivatives of the implicit function.

$$\frac{dh_{t+1}}{dI_{t+1}} = -\frac{\partial g/\partial I_{t+1}}{\partial g/\partial h_{t+1}} \tag{2.22}$$

In our case eq.2.22 is positive so the investment in education increase and human capital increases as crime increases.
To study how crime varies with human capital accumulation the implicit function, using eq.2.13, will be $\phi(h_{t+1}, I_{t+1}) = 0$:

$$\phi(h_{t+1}, I_{t+1}) = I_{t+1} - \delta R S_t - h_{t+1} \tag{2.23}$$

Following the implicit function theorem I can determine how crime varies as a function of human capital accumulation.

$$\frac{dI_{t+1}}{dh_{t+1}} = -\frac{\partial \phi/\partial h_{t+1}}{\partial \phi/\partial I_{t+1}} \tag{2.24}$$

In our case eq.2.24 is negative so as human capital increases, returns from legal activities are higher, the opportunity cost of crime is higher, and there is less incentive to commit crime. Therefore the two functions go in opposite directions and will cross defying the equilibrium level for human capital and crime.

Appendix B

The portfolio choice $h(I_{t+1}, \psi, \delta, \rho)$ is increasing in δ and ψ and decreasing in ρ as:

$$\frac{\partial h(I_{t+1}, \psi, \delta, \rho)}{\partial \psi} > 0 \tag{2.25}$$

$$\frac{\partial h(I_{t+1}, \psi, \delta, \rho)}{\partial \delta} > 0 \tag{2.26}$$

$$\frac{\partial h(I_{t+1}, \psi, \delta, \rho)}{\partial \rho} > 0 \tag{2.27}$$

The occupation choice $I(h_{t+1}, \delta, \rho)$ is increasing in δ.

Appendix C

Education spending is calculated as:

$$\rho\epsilon_t = \rho\frac{\gamma\theta}{\rho R(1-\delta I_{t+1})} = \frac{\gamma\theta}{R(1-\delta I_{t+1})} \tag{2.28}$$

Therefore the share of education spending over GDP is

$$Share_t = \frac{\rho\epsilon_t}{GDP_t} = \frac{\gamma\theta}{R(1-\delta I_{t+1})} \div GDP_t \tag{2.29}$$

There are two methods to calculate GDP at aggregate level. Either I use the spending definition or the income one. GDP calculated with the spending definition is obtained from the sum of education expenditure and savings, as there is no actual consumption in the model:

$$GDP_t = \rho\epsilon_t + S_t \tag{2.30}$$

with

$$S_t = I_t(g_t - \rho\epsilon_t + \omega) - \int_0^{\tilde{\eta}} \eta_i G(\eta_i) d\eta_i + (1-I_t)(h_t - \rho\epsilon_t + \omega) \tag{2.31}$$

Therefore I get

$$GDP_t = I_t\delta R S_{t-1} + \omega + h_t(1-I_t) - \frac{I_t^2}{2} \tag{2.32}$$

The same result is obtained with the income method, where I sum the income of criminals and honest.

$$GDP_t = (1 - I_t)(h_t + \omega) + I_t(g_t + \omega) - \frac{I_t^2}{2} \tag{2.33}$$

where

$$g_t = \delta RS_{t-1} \tag{2.34}$$

$$GDP_t = I_t \delta RS_{t-1} + \omega + h_t(1 - I_t) - \frac{I_t^2}{2} \tag{2.35}$$

The following table shows the calibration and simulation of the model for different values of γ as a check of their sensitiveness to the choice of the parameter γ.

Table 2.6: Education investment over GDP share for different values of γ

Period	γ	ρ	ω	δ	ϵ_t	hc_t	I_t	s_{t-1}	GDP_t	$share_t$
0	30	20	11,80	4,06	0,64	1,90	0,01	0,38	13,70	0,91
1	30	20	11,80	4,06	0,74	3,10	0,01	0,64	14,90	0,85
2	30	20	11,80	4,06	0,76	3,50	0,03	0,72	15,30	0,84
3	30	20	11,80	4,06	0,77	3,60	0,04	0,73	15,40	0,83
4	30	20	11,80	4,06	0,77	3,60	0,04	0,73	15,40	0,83
5	30	20	11,80	4,06	—	3,60	0,04	0,73	15,40	—
0	60	40	24,06	4,06	0,64	1,90	0,10	0,38	25,95	0,96
1	60	40	24,06	4,06	0,69	3,10	0,01	0,64	27,16	0,92
2	60	40	24,06	4,06	0,71	3,40	0,02	0,70	27,46	0,92
3	60	40	24,06	4,06	0,71	3,50	0,03	0,71	27,56	0,92
4	60	40	24,06	4,06	0,71	3,50	0,03	0,72	27,56	0,92
5	60	40	24,06	4,06	—	3,50	0,03	0,72	27,56	—
0	90	61	37,47	4,06	0,64	1,90	0,01	0,38	39,37	0,95
1	90	61	37,47	4,06	0,67	3,10	0,01	0,64	40,57	0,93
2	90	61	37,47	4,06	0,68	3,40	0,02	0,69	40,87	0,92
3	90	61	37,47	4,06	0,69	3,40	0,02	0,70	40,87	0,92
4	90	61	37,47	4,06	0,69	3,40	0,02	0,70	40,87	0,92
5	90	61	37,47	4,06	—	3,40	0,02	0,70	40,87	—

Appendix D

Distribution across districts of the main calibrated parameters.

Figure 2.22: Kernel Density estimate for parameter ρ

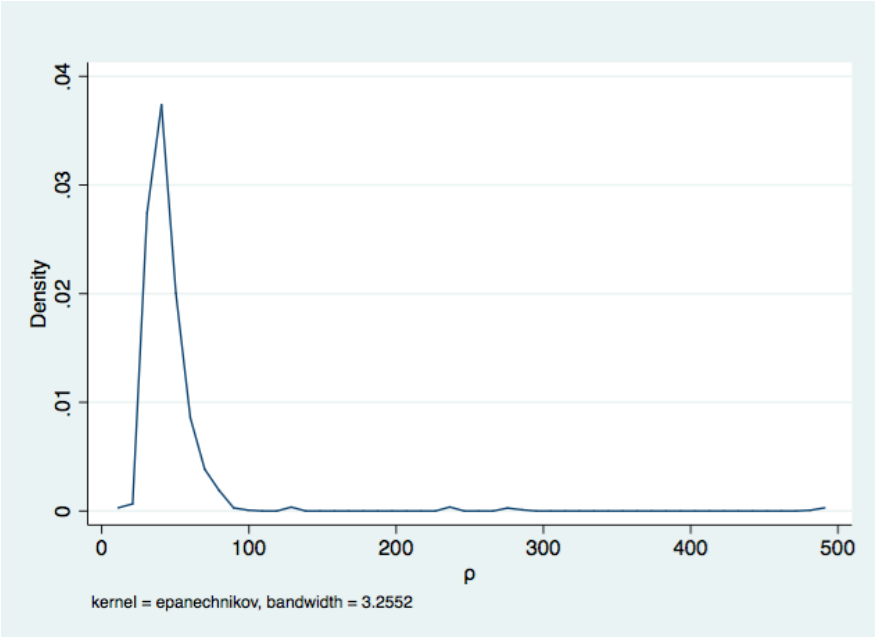


Figure 2.23: Kernel Density estimate for parameter ψ

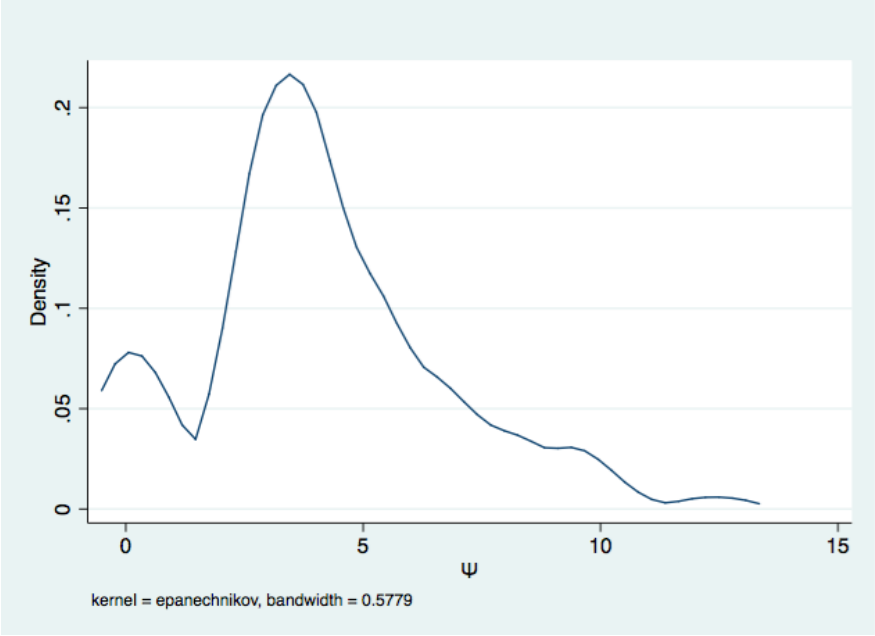
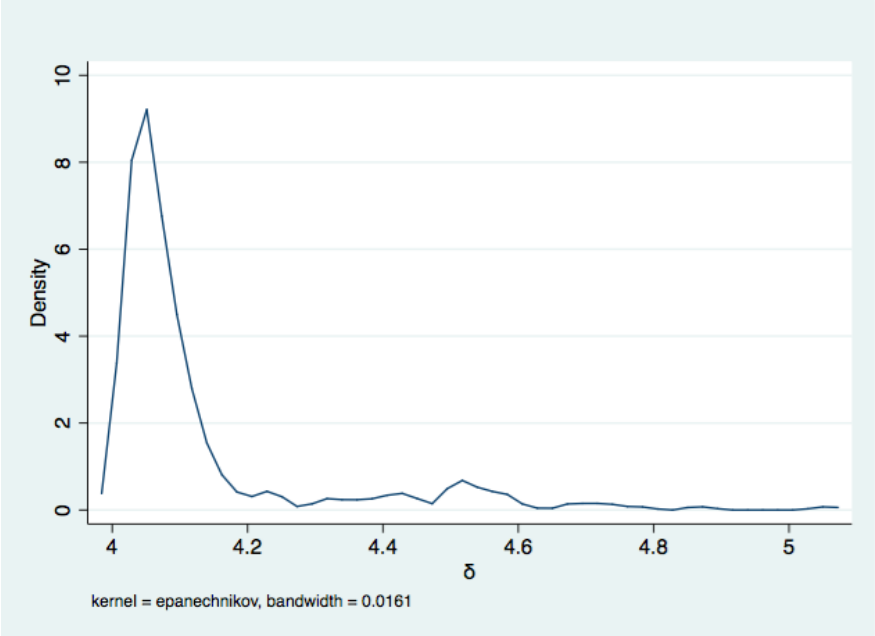


Figure 2.24: Kernel Density estimate for parameter δ



3 Female Education and Fertility in Sub-Saharan Africa: a reappraisal

Co-authored with Renzo Derosas (University of Venice Cà Foscari)

Abstract

In this paper, we study the relationship between female education and births spacing in thirty countries of Sub-Saharan Africa, with the peculiarity of making a more intense use of the information available. Using data from the latest Demographic and Health Surveys, we reconstruct the life histories of mothers and use Cox regressions to analyse inter-birth intervals in the last 5 years before the survey, for higher-order births. Our base model includes covariates at the individual, family, household, community, and country levels. We find that education remains a strong determinant of births spacing, even after controlling for all the other covariates. Furthermore, we explore interactions between education and all the other covariates in the base model, in order to test for significant variations in the effect of education according to different conditions. We find that most of the interactions are significant, supporting the complex and variable links between education and the other factors affecting reproductive behavior. In particular the fertility gap, that is the difference in the risk of having a new pregnancy between educated and uneducated women, increases with age, female autonomy, and the use of contraception. It decreases in the intermediate stages of the Demographic Transition or when the last-born dies. Finally, we find that large differences among countries remain, requiring further investigation.

3.1 Introduction

The relationship between female education and fertility is one of the most investigated topics in the demographic and in the economic literature (Cochrane, 1979; Bledsoe et al., 1999).

Many authors study the education-fertility dyad to define the nature of the relationship, despite underlying complexities and difficulties. Even assuming that the correlation between female education and fertility is not a statistical artefact, questions remain over the pathways through which education exerts its influence (Lloyd et al., 2000; Canning et al., 2015; Bongaarts, 2017), as well as over the need to define the mechanisms through which education affects fertility (Basu, 2002; Casterline et El-Zeini, 2014; Behrman, 2015).

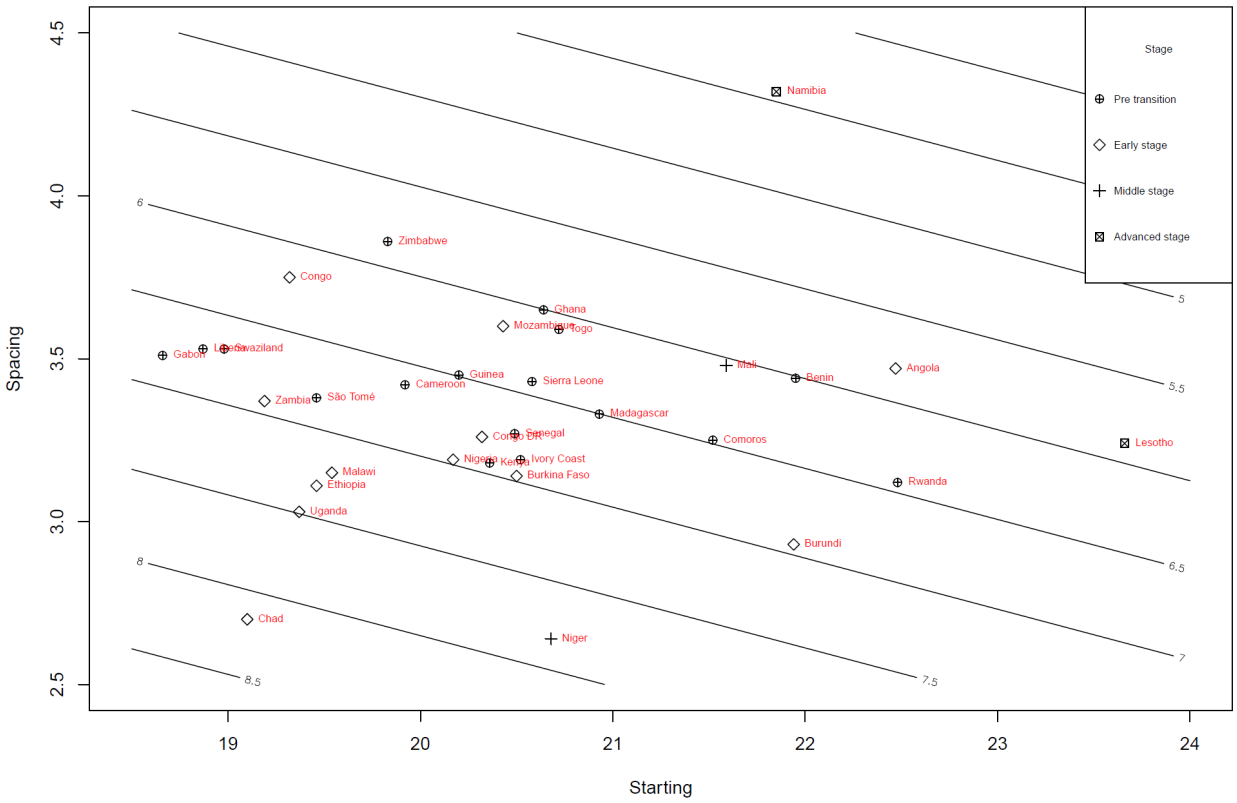
In this paper, we analyze thirty countries of Sub-Saharan Africa to probe the effect of female education on marital fertility, after controlling for, and interacting with, a large number of other factors that also influence reproduction. More specifically, our focus is not on the number of births, but on the length of inter-birth intervals, what demographers refer to as spacing. Spacing is a key determinant of fertility, together with age at first birth, or starting, and age at last birth, or stopping. We ask whether educated women are more likely to adopt longer inter-birth intervals, with beneficial effects on child and maternal well-being (Winikoff, 1983; Conde-Agudelo et al., 2012; Myrskylä and Barclay, 2017) and, ultimately, on the number of pregnancies.

The role of spacing in fertility transitions, past and current, is controversial. Until recently, the received wisdom was that the decline of marital fertility in Western Europe was achieved by shortening the reproductive life span through anticipated, parity-dependent stopping; as for spacing, there is no evidence of a deliberate lengthening of inter-birth intervals (Knodel et van de Walle, 1979; Knodel, 1987; Watkins, 1987). If anything, intervals tended to shorten; when prolonged, it was to postpone a new birth, rather than to limit family size (Van Bavel et Kok, 2004). However, this opinion has been increasingly challenged as factually and methodologically flawed. Critics argue that the distinction between spacing and stopping is a blurred, and that disentangling their specific contribution to fertility limitation is far from straightforward (Anderton et Bean, 1985; Okun, 1995; Hionidou,1998; Friedlander et al., 1999; Fisher, 2000; Van Bavel, 2004a; Van Bavel, 2004b; Cinnirella, Klemp, Weisdorf, 2012, 2017). Whatever the European case, there is growing evidence that contemporary fertility transitions may follow different patterns from those that characterized the historical experience of Western Europe. Whereas, for instance, Taiwan's transition replicated the Western-European pattern, with increased age at marriage and early stopping at low parities, recent studies showed that the fertility drop achieved by several African countries is mostly due to longer birth intervals, while showing no evidence of early stopping (Freedman et al., 1994; Moultrie et al., 2012; Alter, 2016; Towriss et Timaeus, 2017).

This finding is somewhat unexpected. In fact, birth intervals in Sub-Saharan Africa are already quite long, thanks to prolonged breastfeeding and traditional postpartum abstinence

(Caldwell et Caldwell, 1987; Caldwell et al., 1992). The typical age pattern of fertility is almost flat, with early marriage, late age at last birth, and long birth intervals in between (Bongaarts, 2003). This should leave little space for a fertility decline through a further spacing (Bongaarts et Casterline, 2013; Casterline, 2010). Nevertheless, a quick look at aggregate data reveals larger variations in inter-birth intervals than one would expect, with significant impact on a country's fertility. Fig.3.1 displays the average number of children ever born (CEB), or completed marital fertility, in the reproductive space defined by starting and spacing (McDonald, 1984). Countries with similar starting age like, say, Chad and Congo, Niger and Ghana, Burundi and Namibia, differ significantly both in spacing and completed fertility. It is controversial whether such longer intervals are due to the need to postpone births to better times, or to the outright intention to limit the family size (Moultrie et al., 2012; Towriss et Timaeus, 2017). Still, the flexibility of inter-birth intervals turns out to be a key factor for fertility decline in Sub-Saharan Africa and a topic requiring thorough investigation. Our contribution focuses on the effect of female education on the length of birth interval.

Figure 3.1: Age at First Birth (Starting), Interbirth Interval (Spacing) and completed marital fertility in Sub-Saharan Africa



Using data from the latest Demographic and Health Surveys (DHSs), we reconstruct the life histories of mothers and use Cox regressions to analyse inter-birth intervals in the last 5 years before the survey, for higher-order births. Although other scholars followed a similar approach, our analysis makes a more intensive use of the available information. Firstly, we use a Cox proportional hazards model with continuous-time and with time-dependent covariates, as opposed to the more usual discrete-time approach, as a more proper and precise way to deal with time to event analysis (Hosmer and Lemeshow, 1999). Secondly, we adopt a frailty model with random effects at the area unit level, to take into consideration the unobserved heterogeneity at the community level (Kravdal, 2002, 2012). Thirdly, we include in the models two covariates obtained from correspondence analysis and non-hierarchical clustering. Correspondence analysis makes it possible to identify the most relevant characteristics of the household where the woman lives in terms of region, wealth level, media exposure and house facilities. Then, by means of non-hierarchical clustering, we identify four clusters of families with similar characteristics. We use the same approach as far as the area units are concerned. In this way we deal with both the large number of variables provided by DHSs and their

correlation (see Appendix 1 for a description and the results of the procedure). Finally, even if education is given once for all during the early life of the women, most probably before having children, its effect over their reproductive life changes in time, especially through the interactions with other variables. For this reason, we interact our covariate of interest with all the other covariates in the model (except countries) in order to study if there is a significant variation in the effect of education according to different conditions in the covariates.

We find that education remains a strong determinant of spacing, even after controlling for all other covariates, and that significant interactions exist, confirming the complex and variable links between education and the other factors affecting reproductive behavior. In particular, the fertility gap, that is the difference in the risk of having a new pregnancy between educated and uneducated women, increases with age, female autonomy, and the use of contraception. It decreases when the last-born dies or in the intermediate stages of the Demographic Transition. Furthermore, we find that large differences among countries remain, requiring further investigation.

The study of the relationship between education and spacing is part of the most extensive and articulated literature dealing with the relationship between education and fertility. Since the mid-1970s, the relationship between female education and fertility is one of the most investigated topics (Cochrane, 1979; Bledsoe et al., 1999). Four decades of studies gathered an ever-growing mass of evidence highlighting the importance of education, and in particular of female education, for fertility decline, past and present, with few and non-conclusive exceptions (Cochrane, 1979; Sandhu, 1996; Diamond et al., 1999; Lloyd et al., 2000; Monstad et al., 2008; Bongaarts, 2010; Canning et al., 2015; Vogl , 2015; Tertilt et al., 2016; NAS, 2017). However, the expansion of the empirical evidence went hand in hand with an increasing awareness of the complexities and difficulties lying behind the education-fertility dyad. These complexities include the definition and measurement of education (Carter, 1999), the functional form of the relationship, and its dependence on and interaction with a variety of contextual factors (Lesthaeghe et al., 1985; Castro Martin, 1995; Jejeebhoy, 1995; Eloundou-Enyegue, 1999; Diamond et al., 1999; Garenne, 2012; Behrman, 2015; De la Croix and Gobbi, 2014; NAS, 2017; Towriss et Timaeus, 2017).

The nature of the relationship between education and fertility was also put under scrutiny. In particular, scholars argued that an observed negative correlation might derive from heterogeneity, reverse causation or endogenous association rather than from causation (Wojtkiewicz and McDonald, 1987; Lloyd, 1994; Eloundou-Enyegue, 1999; Dasgupta, 1999; Drèze, 2001; Angeles, 2005; Behrman, 2015; Bongaarts, 2017).

Even assuming that the correlation between female education and fertility is not a statistical

artefact, questions remain of the pathways through which education exerts its influence (Lloyd et al., 2000; Canning et al., 2015; Bongaarts, 2017). This is not an easy task to carry out since the link between education and fertility is not straight nor simple (Bulatao and Lee, 1983; Mason, 1984; Castro Martin and Juàrez, 1995; United Nations, 1995; Diamond et al., 1999; Behrman, 2015). However, although the efforts to identify the proximate determinants are useful, they do not replace the need to define the mechanisms through which education affects fertility (Akman, 2002; Aurig, 2013; McDevitt et Johnson, 2013; Shapiro, 2015; De la Croix et Gobbi, 2017).

Among the many authors investigating these mechanisms (Basu, 2002; Casterline et El-Zeini, 2014; Behrman, 2015) Diamond et al. (1999) identify five major classes: key contextual factors (Kravdal, 2002,2012; Schoumaker et al., 2006; Derose et Kravdal, 2007); the skills and knowledge imparted by schooling; social and ideational influence (Caldwell, 1982; Castro Martin and Juàrez, 1995; Basu, 2002; Akman, 2002; Tertilt et Doepke, 2018); autonomy; and enhanced employment opportunities for the educated (Caldwell, 1982; Dyson and Murphy 1985; Garenne, 2012, 2014; De la Croix and Gobbi, 2014).

The work of these authors guided our choice of the variables to include in the model and to interact with the main explicative variable, education. As a result we are able to open the black box of the education-fertility dyad, with a precise idea of the direction where to point our analysis and with the aim to study more in depth the relationship between education and birth intervals, as a way to better understand the global relationship between education and fertility.

3.2 DHS Data for Longitudinal Analysis

3.2.1 Data and Methods

Our data is based on the latest DHSs available containing information about the health and nutritional status of several households (from 5,000 to 30,000) for each country in the region of interest as well as on educational and demographic conditions. The DHSs sample households in interview groups on a regional base in order to obtain a wide representation of all the regions in the countries, from both rural and urban areas. Each region is divided in clusters or CEAs (Census Enumeration Areas). A CEA covers a territory including either a few villages or a small town or portion of a big town (Kravdal, 2012). Each CEA includes around 1000 individuals of which 250 are women aged between 15 and 49, that

is fertile women eligible for individual interviews. Among all the CEAs in a region, some are randomly selected for the survey. For each selected cluster, 30 households are chosen with systematic sampling. Families members are then evaluated to decide whether they are eligible for the individual survey. This sampling process is useful to account for regional and rural and urban differences and to obtain countries' samples with an adequate number of qualified women and men for the individual interview.

Given the large amount of information and the structure of the data contained in the DHSs, we decide to base our work on the event history analysis method, which defines the entire reproductive history of women (date of first birth, dates of successive births, birth intervals, child death, abstinence, amenorrhea, sterilization,...) and studies the role of different variables on the single events forming the reproductive history, with particular attention to births (Tsuya et al, 2010). In the specific we build our database using a longitudinal perspective. Each woman, aged between 15 and 49, is associated with a time-line accounting for the births and the deaths of children and other relevant occurrences. Observations are right-censored when women reach the age of 50, remain widow, undergo sterilization or at the interview date. Generally, in survival analysis the event of interest is a transition from one status to another, in our case the transition from one parity to another. The delivery intervals measure the duration of the transition, or the spacing among births. For each woman we have different inter-birth spells, obtained using the information on the birth dates of her children. For any two births, where the most recent one is the birth of interest, a spell begins 30 days after the preceding birth, assuming that it takes one month for the ovulation to restart [1], and ends 270 days before the birth of interest, considering that during the pregnancy there is no ovulation. If the inter-birth spell ends due to a pregnancy the dummy variable "event" associated to that spell will have value "1". However, a period can end without the occurrence of a conception, but for a change in time dependent variables (age of the mother, net parity or the stage of the demographic transition) associated to each woman. If the inter-birth spell ends due to one of these changes and not for a new birth, the dummy variable "event" associated to that spell will have value "0". For reasons explained below, we restrict the time frame of the analysis to intervals occurred in the five years before the survey. Moreover, even if country's weights are available in the database, we do not weight the observations. Our approach is a micro one, we are interested in the effect that the variables have on the individual decisions, not on the entire population.

Using this database we model the extension of delivery intervals. That is we define the

[1] We are assuming that it takes 1 month for the ovulation to restart, even if this would normally required 3 months. However this would mean loosing too many intervals and information, such that we decide to use the assumption of 1 month.

effects, or relative risks, of education on birth intervals, to test if the impact of education remains relevant even after controlling for a large number of other factors which also influence fertility. We use a Cox proportional hazards model with continuous time and time-dependent covariates and, to take into consideration the unobserved heterogeneity at community level, we implement a frailty model with random effects at area unit rank. Our base model includes covariates at the individual, family, household, community and country levels. In addition to this, we interact our covariate of interest with all the other covariates (except countries), to study if there is a significant variation in the effect of education according to different conditions in the other variables.

3.2.2 Potential Problems

The use of DHS data for event history analysis is not without concerns. In this paragraph we present the main potential problems to be discussed for a more informed interpretation of the results.

First and foremost, issues have been raised on the quality of DHS data in Sub-Saharan Africa. The main problem is the misreporting of births, which can result from different misleading behaviors of the women or the interviewers. For example, children's birth or ages can be heaped. Moreover, as for children younger than 5 there is an additional questionnaire, in many cases the interviewer reports the child to be older, to avoid asking more questions. Indeed, there is an irregular fertility trend with a decrease in TFR the 5th year before the survey and an increase in the 6th year. Women may also tend to report distant births as closer to the moment of the interview (Potter Effect). Finally some women just omit recent births, especially if children are deceased, to avoid additional or sensitive questions (Kravdal, 2002; Fuchs, 2011; Garenne, 2012; Schoumaker, 2014; Towriss and Timaeus, 2017).

Another problem may arise from the fact that many of the explicative and control variables of the model describe the situation of the woman at the time of the interview or in the few years before. It can be questionable to study the effect of these variables on births occurred almost 20 years before.

For both these reasons we restrict the time frame of the analysis to the intervals occurred in the five years before the survey, even if the period of analysis is usually restricted to three years. In order to avoid a loss of valuable information we decide to use the larger, and still safe, interval of 5 years (tab.3.1). For robustness check, Appendix 2 shows the results of the base model when using a database including intervals up to 10 years, 3 years and 5 years before the survey. Using the 10 years database allows for smaller standard errors and better

p-values but the uncertainty associated with the quality of data and the problems regarding conducting an analysis over such a long period of time, prevent us from using this database. Considering that results do not change significantly in the 3 or 5 years interval database, we decide to use the 5 years one. This choice yields to 355840 intervals used to model the risk of higher order births. In the base model 88587 events, or conceptions, occur. Tab.3.1 shows, for each country in the study, the time span considered according to the year of the latest available survey, the average Total Fertility Rate in the country, the average inter-birth interval in years, the number of intervals and the number of events included in the analysis.

Table 3.1: Surveys' Summary Information by Country: time span covered by the survey for each country, Total Fertility Rate for 2015, average spacing in years, number of intervals and number of events occurred

Country	Time Span	TFR 2015	Birth Interval (Y)	Number of Intervals	Number of Events
Benin	2006 – 2011	5	3	15472	3932
Burkina Faso	2005 – 2010	5	3	18581	4379
Burundi	2005 – 2010	6	2,6	8774	2721
Cameroon	2006 – 2011	5	3	13418	3690
Chad	2007 – 2012	6	2,6	6688	2061
Comoros	1999 – 2004	4	2,9	3453	1067
Congo, Dem. Rep.	2008 – 2013	6	3	20768	6958
Congo, Rep.	2006 – 2011	5	3,4	10987	2562
Cote d'Ivoire	2007 – 2012	5	3,1	8454	2084
Ethiopia	1998 – 2003	4	2,8	17183	3709
Gabon	2007 – 2012	4	3,1	6072	1568
Ghana	2003 – 2008	4	3,3	3428	766
Guinea	2007 – 2012	5	3,2	7627	1874
Kenya	2003 – 2008	4	3	7029	1871
Lesotho	2004 – 2009	3	3,7	4835	740
Liberia	2008 – 2013	5	3,1	8228	1948
Madagascar	2003 – 2008	4	2,9	15371	3704
Malawi	2005 – 2010	5	3	25994	5742
Mali	2007 – 2012	6	3,1	14378	3419
Mozambique	2006 – 2011	5	3,3	12727	3172
Namibia	2008 – 2013	3	3,8	3922	621
Nigeria	2008 – 2013	6	2,9	36433	10680
Rwanda	2005 – 2010	4	2,9	14938	1493
Senegal	2005 – 2010	5	3	13916	3890
Sierra Leone	2008 – 2013	5	3,1	12969	3050
Swaziland	2001 – 2006	3	3,1	2559	461
Togo	2008 – 2013	5	3,2	8175	1800
Uganda	2006 – 2011	6	2,7	10408	2818
Zambia	2008 – 2013	5	3	16263	3777
Zimbabwe	2005 – 2010	4	3,6	6790	1066

Source: STATcomplier and own estimations

Further issues may arise if the key assumptions of the Cox model are violated, endangering the validity of the model. The first of these assumptions is the non-informative character of the censoring, determining that observations must not be excluded because of the probability

of an event. In our case the censoring occurs when women reach a certain age or at the interview date, therefore the mechanism of censoring is not related to the probability of a conception. The second key assumption is the one regarding the proportionality of the hazard ratio. When a variable has more strata, the hazard function of each category must be proportional over time, that is the relative hazard must be constant (Park and Hendry, 2015). We checked this assumption with the Grambsch-Therneau test on the Schoenfeld residuals (Appendix 3). From appendix 3 we can conclude that the proportionality assumption holds.

Another problem we needed to take into account is the correlation among the explanatory variables. In this case we use the variance inflation factor (VIF) to determine the optimal group of explanatory variables. The common process is to evaluate the VIF for a proposed set of variables and to exclude from the model the variables with a factor higher than a threshold which may vary from 5 to 10. By following this method we select the list of covariates in tab.3.2 in the following section. A potential problem remains for the variable "Stage of Demographic Transition", which is highly correlated with the country covariate as in each country we have at most two different stages (Appendix 4). However, we cannot exclude the dummy of the country of origin of the mother from the model as it is necessary to control for her background at a macro level, as well as we cannot exclude the variable concerning the stage of the Demographic Transition because of its interaction with education. Appendix 5 shows the base model without the variable "Stage of Demographic Transition". Excluding one of the two variables does not alter the main conclusions of our analysis. We decide therefore to keep both without ignoring however the potential correlation.

3.2.3 Variables Selection

Our base model includes covariates at different levels: individual, family, household, area unit, and country, as in tab.3.2.

Table 3.2: Variable Selection in the Base Model

Level	Covariates
Individual	age, use of birth-control methods, age at first intercourse relation to household head, years of education
Family	number of surviving children, survival of the prevoius born length of cohabitation, husband's years of education, difference with husband age
Household	classification from cluster analysis (urban or rural, quality of housing water availability, access to media, property of means of transport), religion
Area Unit	classification from cluster analysis area unit average female education
Country	stage of demographic transition, a dummy for countries

The principal independent variable of the model is education measured in years at individual and area level. In the DHS surveys, data on education consist of four variables. "Higher education level" divides respondents into no education, primary, secondary or higher. "Educational attainment" recodes the highest level of education completed. "Highest year of education" counts the number of years a person studied. Information is also given on the literacy rate distinguishing among people able to read part of a sentence, a complete sentence or unable to read at all. Our analysis uses the number of years spent on education and the average length of education at the area unit level, to measure the effect of the community level of education on the behaviour of the individual. Kravdal (2002, 2012) has already stressed the importance of this component of education, showing that community education is relevant independently of individual education. To conclude, there is no information on the type of education received by women, even if this would have been a valuable point to add to the analysis (Oye, Pritchett, Sandefur, 2018).

In addition to female education variables the base model includes some control variables. Age of the mother at birth, age at first sex, use of birth control, number of surviving children, parity and religion describe the family framework where decisions about fertility are taken. Relation to the household head, survival of the previous child, difference with husband age and husband's years of education refer to the role of the woman in the family, her autonomy

and her bargaining power when taking important decisions on key matters such as fertility. The stage of Demographic Transition takes into account the fact that in many regions of Africa the process toward a modern demographic regime is not complete. This variable is built starting from the average TFR of a country at rural and urban level, considering that the fertility decrease might start at different moments and proceed with different paces in urban versus rural area (Casterline, 2001). The categories of this variable, both for women living in rural or urban area, were determined using the classification of Gebreselassie (2011), that is "Pre-transition" if the TFR is higher than 7, "early transition" if the TFR is between 6.9 and 5 children per woman, "Middle Transition" if TFR is between 4.9 and 3, and "Late/Post Transition" if women have between 3 to 2.1 children. Finally two variables obtained from correspondence analysis and non-hierarchical clustering are added to the model, one at the household level and one at the area level. Each of these variables has four categories which represent four classes with specific characteristics in terms of residence, facilities, means of transport, wealth level and access to media. Tab.3.3 provides some summary information on the classes, however in Appendix 1 the detailed profiles of the clusters at household and area level are described in tab.3.6 and tab.3.7.

Table 3.3: Classes Profile at Household and Area Unit Level

Level: Household	Classes Caracteristics
Class 1	rural, no facilities (pipes, toilet,electricity), wood/earth floor, water from spring/bottle cane wall, no radio, no tv, never listen to radio/tv, agricultural sector, african religion, poor
Class 2	mostly rural, some facilities (pipes, toilet, no electricity), no tv, yes radio, yes bike, yes scooter, metal floor, bricks walls, unemployed or manual work, middle class
Class 3	urban, yes facilities (electricity,...), yes tv, yes fridge, cement floor, read newspaper, listen to radio, watch tv often, manual or service sector, rich
Class 4	urban, yes facilities (electricity,...), yes tv, yes fridge, cement floor, read newspaper, listen to radio, watch tv often, manual or service sector, protestant, rich
Level. Area Unit	Classes Caracteristics
Class 1	rural, no facilities (pipes, toilet,electricity), wood/earth floor, water from sping/bottle cane wall, no radio, no tv, never listen to radio/tv, poor
Class 2	urban, yes facilities (pipes, toilet, electricity), tv, fridge, car, cement floor and walls, read newspaper, listen to radio, watch tv once a week, rich
Class 3	urban, yes facilities (pipes, toilet, electricity), tv, fridge, car, cement floor and walls, read newspaper, listen to radio and watch tv every day, rich
Class 4	rural, no facilities (pipes, toilet,electricity), wood/earth floor, water from sping/bottle cane wall, yes bike, yes scooter, no radio, no tv, never listen to radio/tv, middle class

3.3 Results

Tab.3.4 shows the results of the Cox regression while tab.3.5 shows the outcome of the set of interactions between education and the principal explicative variables. For each of these variables we use the base model adding the interaction of a covariate with education, one at a time. The hazards ratio of tab.3.5 are obtained multiplying the hazards ratio of the two variables with the one of the interaction in the interaction augmented models. For a an easier interpretation, in tab.3.5 we add a column showing the proportion of the highest to the lowest educational level hazards ratio: the further the value is from 1, in a positive or negative direction, the stronger is the effect of the interaction and the greatest is the difference in the behavior between educated and uneducated women. Finally, we consider an interaction significant only if it adds enough precision to the model. To measure this, we calculate two time the difference in the log likelihood of the two models (the one with the interaction and the base one). If the difference is big enough, then the interaction model is significant and can be added in the analysis [2].

First and foremost, the impact of education on births spacing remains relevant, even after controlling for a large number of other factors which also influence fertility. The difference in the hazards ratio between the most and the less educated women is as large as 13 per cent. Interestingly, until 6 years of schooling there is no significant difference with the uneducated group, coherently with the threshold effect mentioned in the literature (Bledsoe et al., 1999; De la Croix and Gobbi, 2014). Husband's education levels influence reproductive decisions in a less decisive way (Clealand, 2002), as the difference in the risk for women with a more educated husband versus women with a non educated spouse is "only" 8 per cent. Also, we did not find any evidence of an additional community effect of female education beyond what is accounted for using a frailty model at area level.

The relationship between education and birth spacing can be better understood looking at the interactions of education with the other covariates (tab.3.5).

The interaction between the use of birth control and education supports the claim that education is a key element to improve and extend the use of contraception (Brass, 1993, Casterline, 2001). The difference in the risk of a pregnancy between educated and uneducated

[2]Large enough means that the difference in the two log-likelihood is higher than the value of the χ^2 distribution with the same significance level of the model and a number of degrees of freedom equals to the difference of degrees of freedom of the two models. For example: If the difference in the log-likelihood of the two models is 30 and the difference in the degrees of freedom of the two models is 7, the value of the χ^2 distribution corresponding to the significance level of 5% and to 7 degrees of freedom is 14.09. Therefore as 30 is higher than 14.09, we can conclude that the interaction is adding precision to the model and therefore is considered significant.

women is only 13 per cent when no methods are used, it becomes 89 per cent when such practice is used. The effectiveness of the contraceptive methods is incomparably greater for educated women than for uneducated ones. Some may claim that this result is driven by the fact that more educated women are often using more modern and more effective methods of contraception. This is a plausible explanation, which deserves more attention. However, for now our focus is on the multiplicative effect of education on the use of birth control, for any type of method used.

The interaction between the age of the mother and education shows that the effect of education is much larger at a late age than in the 20s and 30s. In other words, educated women tend to reduce the period of reproductive activity with respect to uneducated ones. This is a fundamental result in a region like Sub-Saharan Africa, which is characterized by a flat age pattern of fertility with exceptionally late stopping (Bongaarts, 2003; Bongaarts and Casterline, 2013). Education is therefore associated with a demographic change towards a more regular, and predictable, fertility pattern where the greatest fertility decline occurs at older ages (Cohen, 1998; Fenn, 2013; Bongaarts, 2017).

Also women empowerment is a key determinant of fertility (Clealand, 2002). Female autonomy makes women more active in taking decisions, especially on fertility (Jejeebhoy, 1995). In households where women have a good status, they have more bargaining power, there is more gender equality and communication between spouses, such that women preferences on fertility may prevail (Diamond et al., 1999; Akman, 2002; Basu, 2002; Clealand, 2002; Bankole, 2011; Canning et al., 2015). In the base model the variables "Difference with husband's age", "Relation to head", "Survival of previous child" and "Husband years of education", are the proxy of women empowerment.

The interactions of education with these variables show that when women are close to the age of their husband (categories "$0-4$" or "older") and they are more educated, their hazards ratio diverges from the one of less educated women in favour of a fertility reduction. The same is true for the role of the women in the family. When the woman is the head of the household and she has completed more years of school, the risk of a new birth drops with respect to less educated women with the same role in the family or to women with less powerful positions (wife or daughter). Finally when the previous born dies the chances of a new pregnancy are greater for educated women rather than for uneducated mothers. An outcome that suggest an intentional replacement rather than an accidental event due to the interruption of breastfeeding and the restart of ovulation. Again education has a multiplicative effect for female empowerment.

The last significant interaction is the one of education with the stage of Demographic Transi-

tion. Many authors claim that the relationship between education and fertility is unequivocal during the transition, while the same is not true before and after the process (Castro Martin, 1995; Clealand, 2002; Garenne, 2012). In particular Clealand (2002) claims that in a pre-transitional or post-transitional period the tie between education and fertility is weak and diversified, while during the transition fertility differentials by schooling mitigate until convergence.

The results endorse the thesis of Clealand (2002). In the early or late stages of the process women with different levels of education behave in different ways. In the middle of the transition, education becomes a driver of transformation at any level, regardless of the years of schooling. The effect of education on fertility is therefore amplified during the central stages of the Demographic Transition.

In addition to interactions, the cluster variables for household and area unit are another fundamental element of the model. The detailed description of each categories of these two variables is given in tab.3.3 and in Appendix 1.

As we can see from these tables socio-economic conditions differ among clusters. These differences are often mapped into different fertility outcomes (Fuchs, 2011). To cite a few examples, the onset of the fertility decline or the pace and the pattern of this decline differs between rural and urban areas (Casterline, 2001; Lesthaeghe, 2014; Shapiro and Tenikue, 2015); also, different wealth levels have different effects on fertility (Benefo, 1994). According to Fuchs (2011) it is necessary to disentangle the socio-economic indicators into their many components as they have a different repercussion for fertility outcomes. The results from our analysis support the role of socio-economic factors, different from education, as associated to fertility. Richer clusters at household level, like clusters 3 and 4, with more facilities (car, electricity, bicycle, ...) and more access to media, present lower risk of a new conception in the family. For area unit clusters the same is true. In cluster 3, where most of the households are rich, with intense use of facilities and media, the probability of a new birth is 13 per cent lower than in cluster 1, which is characterised by areas with a majority of poor households with no facilities and no access to media.

Finally, notwithstanding the number of significant controls present in the model, huge differences by countries remains, which can be as large as 100 per cent or more. This outcome deserves further investigations with country specific analysis.

Table 3.4: Cox Regression Results: Effect of the levels of education and of the other covariates on the risk of second and higher order pregnancies (Hazard Ratio), for women aged 15-49, in the 5 years before the survey.

Covariate	*coef*	*exp(coef)*	*se(coef)*
Parity			
1-2	0,000	1,000	*reference*
3-4	-0,090	0,914 ***	0,015
5-6	-0,067	0,935 ***	0,019
7-9	-0,074	0,928 **	0,024
10+	-0,224	0,799 ***	0,039
Survived Children			
0-2	0,000	1,000	*reference*
3-5	-0,107	0,899 ***	0,015
6+	-0,142	0,867 ***	0,022
Previous Child is			
alive or dead since 1+ years	0,000	1,000	*reference*
dead < 1 year	1,124	3,077 ***	0,015
Use of Birth Control			
No	0,000	1,000	*reference*
Yes	-3,376	0,034 ***	0,052
Mother's Age at Childbirth			
20-29	0,000	1,000	*reference*
< 20	0,012	1,013	0,012
30-39	-0,276	0,759 ***	0,010
40+	-1,105	0,331 ***	0,023
Age at First Sex			
< 15	0,000	1,000	*reference*
15-19	0,110	1,116 ***	0,014
20+	-0,043	0,958 **	0,014
Difference with Husband Age			
0-4	0,000	1,000	*reference*
5-9	-0,030	0,971 **	0,009
10+	-0,055	0,947 ***	0,010
< 0	0,018	1,018	0,018
unknown	-0,527	0,590 ***	0,018
Relation to Head			
Wife	0,000	1,000	*reference*
Head	-0,153	0,858 ***	0,010
Daughter	-0,137	0,872 ***	0,026
Other	-0,202	0,817 ***	0,018

Events, n = 88587, 355840

Integrated Log-likelihood = -992235.3

Integrated Log-likelihood Chisq = 38769.84 df 72

Frailty Model (RE) at Area Level

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

Covariate	$coef$	$exp(coef)$	$se(coef)$
Cluster Analysis: Household Level			
Cluster 1	0,000	1,000	reference
Cluster 2	-0,061	0,941 ***	0,009
Cluster 3	-0,129	0,879 ***	0,016
Cluster 4	-0,167	0,847 ***	0,024
Female Education (years)			
0-2	0,000	1,000	reference
3-6	-0,010	0,990	0,010
7-9	-0,056	0,945 ***	0,015
9+	-0,144	0,866 ***	0,016
Male Education (years)			
0-2	0,000	1,000	reference
3-6	0,008	1,008	0,010
7-9	-0,030	0,970 *	0,014
9+	-0,080	0,923 ***	0,012
NA	-0,029	0,971	0,021
Religion			
Muslim	0,000	1,000	reference
Catholic	-0,029	0,971 *	0,014
Protestant	-0,013	0,987	0,015
Other Catholic	-0,019	0,981	0,014
Traditional	0,021	1,021	0,026
Other	0,015	1,015	0,019
Cluster Analysis: Area Level			
Cluster 1	0,000	1,000	reference
Cluster 2	0,028	1,028	0,023
Cluster 3	-0,138	0,871 ***	0,025
Cluster 4	0,016	1,016	0,013
Average Area Female Education	0,003	1,003	0,003
Stage of Demographic Transition			
Pre-Transition	0,000	1,000	reference
Early transition	-0,176	0,838 ***	0,023
Mid Transition	-0,268	0,765 ***	0,027
Late/Post-Transition	-0,498	0,608 ***	0,052

Events, n = 88587, 355840

Integrated Log-likelihood = -992235.3

Integrated Log-likelihood Chisq = 38769.84 df 72

Frailty Model (RE) at Area Level

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

Covariate	$coef$	$exp(coef)$	$se(coef)$
Country			
Benin	0,000	1,000	reference
Burkina Faso	-0,052	0,949	0,028
Burundi	0,477	1,610 ***	0,033
Cameroon	0,452	1,572 ***	0,031
Chad	0,423	1,526 ***	0,037
Comoros	0,553	1,739 ***	0,043
Congo	0,391	1,478 ***	0,035
Congo DR	0,472	1,604 ***	0,034
Ethiopia	0,181	1,198 ***	0,031
Gabon	0,631	1,879 ***	0,040
Ghana	0,085	1,089	0,045
Guinea	-0,125	0,882 ***	0,035
Cote d'Ivoire	0,076	1,079 *	0,034
Kenya	0,482	1,620 ***	0,038
Lesotho	-0,093	0,912	0,050
Liberia	0,100	1,105	0,035
Madagascar	0,249	1,282 ***	0,031
Malawi	0,180	1,197 ***	0,028
Mali	0,068	1,070 *	0,031
Mozambique	0,191	1,210 ***	0,030
Namibia	0,229	1,257 ***	0,053
Nigeria	0,395	1,484 ***	0,026
Rwanda	0,213	1,237 ***	0,034
Senegal	0,255	1,290 ***	0,031
Sierra Leone	0,007	1,008	0,031
Swaziland	0,279	1,322 ***	0,058
Togo	-0,025	0,975	0,035
Uganda	0,611	1,843 ***	0,033
Zambia	0,217	1,243 ***	0,036
Zimbabwe	0,028	1,028	0,046

Events, n = 88587, 355840

Integrated Log-likelihood = -992235.3

Integrated Log-likelihood Chisq = 38769.84 df 72

Frailty Model (RE) at Area Level

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

Table 3.5: Interactions Results: Effect of the interaction of education with the other variables on the the risk of second and higher order pregnancies (Hazard Ratio). Proportion of the highest to the lowest education level hazard ratio for each category of each covariate.

Covariates	Years of Education				
	0-2	3-6	7-9	9+	
Parity					Highest/Lowest
0-2	1,00	1,00	0,95	0,88	0,88
3-4	0,93	0,89	0,93	0,77	0,83
5-6	0,94	0,93	0,88	0,82	0,88
7-9	0,92	0,95	0,94	0,85	0,93
10+	0,77	0,89	0,82	1,03	1,34
Net Parity					
0-2	1,00	1,01	0,96	0,88	0,88
3-5	0,92	0,88	0,84	0,76	0,82
6+	0,86	0,90	0,85	0,79	0,92
Last Child is					
Alive or dead 1+ years	1,00	0,98	0,93	0,85	0,85
Dead < 1 year	2,81	3,12	3,44	3,55	1,27
Use of Birth Control					
No	1,00	1,00	0,95	0,87	0,87
Yes	0,09	0,02	0,01	0,01	0,11
Mother's Age					
20-29	1,00	0,98	0,93	0,87	0,87
< 20	0,97	1,05	1,04	0,83	0,85
30-39	0,76	0,75	0,71	0,65	0,87
40+	0,35	0,31	0,24	0,24	0,67
Difference with Husband's Age					
0-4	1,00	0,96	0,92	0,82	0,82
5-9	0,95	0,94	0,89	0,86	0,90
10+	0,91	0,92	0,93	0,88	0,96
< 0	1,03	0,98	0,86	0,81	0,78
NA	0,60	0,61	0,55	0,40	0,67
Relation to household head					
Wife	1,00	0,99	0,94	0,91	0,91
Head	0,88	0,87	0,82	0,68	0,78
Daughter	0,85	0,86	0,89	0,75	0,88
Other	0,82	0,86	0,82	0,61	0,75
Stage of Demographic Transition					
Pre-Transition	1,00	0,91	0,90	0,78	0,78
Early Transition	0,80	0,81	0,78	0,71	0,89
Mid Transition	0,74	0,73	0,67	0,64	0,86
Late and Post Transition	0,66	0,51	0,53	0,49	0,75

The reference value is the one corresponding to the first line and first column.

Highest/Lowest is the ratio between the hazard ratio of the highest level of education and the hazard ratio of the lowest level of education for each category of the covariates.

3.4 Conclusions

Our analysis of inter-birth intervals and fertility in 30 countries of Sub-Saharan Africa used the latest datasets available from the Demographic and Health Surveys. We adopted a modelling strategy different under many aspects from that used by other scholars who worked in the same field and with the same source material. We used a continuous-time approach, and included new covariates, who turned out to be statistically significant. Finally, we included interactions between education and the other covariates. The results of our analysis stress that education has an important and independent impact on the length of inter-birth intervals, and therefore on fertility, even after controlling for a large set of other variables. In addition to this all the other covariates included in the model affect inter-births intervals and fertility in the expected direction, with the exception of the average education in the area unit, which results to have no additional effect on fertility beyond what is accounted for in the frailty model with random effects at area level. Other covariates at the community level are significant. The regression also reveals an important difference among countries, which the model does not account for, and that requires further investigation. Interactions show several interesting features of the way education influences reproductive behavior in different conditions, mostly increasing the gap in the risk of having a new pregnancy between educated and uneducated women, but in some cases decreasing it. The gap increases with age, female autonomy and use of contraception; it decreases when the last born dies or in the intermediate stages of the Demographic Transition.

3.5 Appendix

Appendix 1: Correspondence Analysis and Non-Hierarchical Clustering

To deal with the large number of variables provided by DHS and to deal with the fact that many of these variables are correlated, we introduce two variables obtained with the correspondence analysis and the non-hierarchical clustering methods.

First, from the correspondence analysis we obtain a set of factors which represent the most frequent combinations of the categories of the original variables. The factors are new quantitative variables which simplify the correlation table between the original covariates. Usually the most relevant factors present characteristics in opposition the one to the other.

Figure 3.2: Factor 1 versus Factor 2

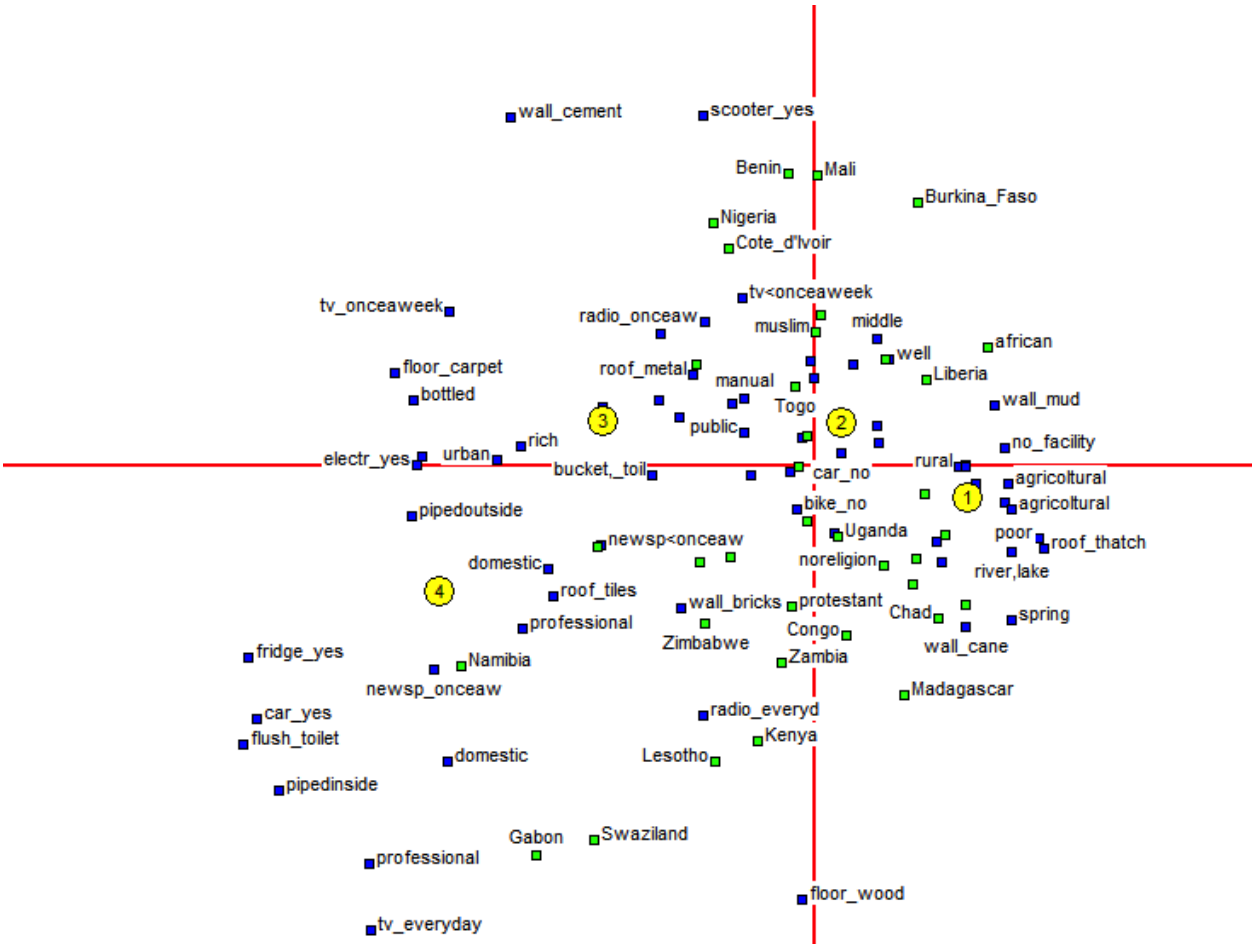


Figure 3.3: Factor 1 versus Factor 3

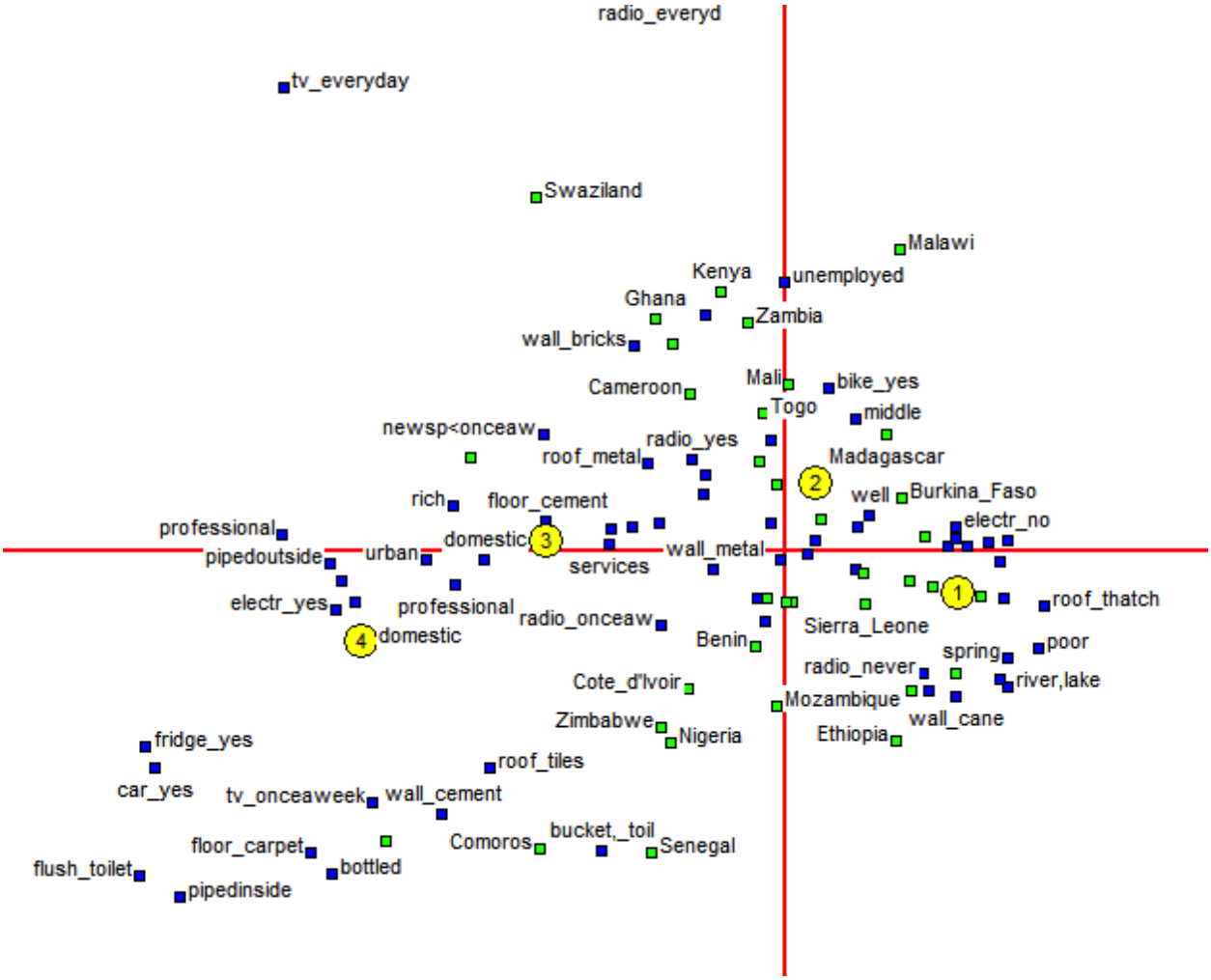


Fig.3.2 and fig.3.3 are two examples of correspondence analysis projections on a chart where we plot factorial axes, or factor, 1 and 2 and then 1 and 3. Along each axes the categories of a variable take opposite positions (urban versus rural, tv yes versus tv no,...).

To proceed, to each observation is associated a value for each relevant factor. This value is a coordinate. The entire string of coordinate for one observation determine the position of that observation in the multidimensional space formed by all the relevant factorial axes. After having determined the position of each element, we are able to create clusters of observations which are close to each other in the multidimensional factor space. Each cluster of elements have some particular characteristics from the socio-economic point of view. We undertake this process for variables describing the characteristics of the household where the woman is living and for variables describing the area, or community, of the woman. In this way we have two set of clusters, one for the household characteristics and one for the area characteristics. These clusters are the categories of the two new variables included in the

model. For obvious reasons we can not give a visual representation of the multidimensional factor space associated to household and areas, however tab.3.6 and tab.3.7 show the profiles of the four main clusters, or classes, according to the intensity of the main factors obtained from the correspondence analysis. The symbols "+" "∼" and "-" are used to describe the profiles of the classes. "+" is associated with a positive presence of a particular factor in that class. Following the same criteria "-" is associated with a negative presence and "∼" with an unbiased presence of a particular factor in a particular class.

Table 3.6: Household Classes Profiles

CLASS	rural	urban	well	piped inside	public	piped outside	bottled	spring	river,lake	flush toilet	pit latrine	bucket
1	94.2	5.8	46.6	0.5	9.4	0.7	0.3	21.5	21.0	0.2	48.4	0.9
	++	—-		—-	–	—-	—-	++	++	—-	–	–
2	78.9	21.1	56.0	1.4	22.4	1.5	0.7	9.7	8.3	0.9	76.6	1.0
		–	++	—-	++	—-	—-	–	–	—-	++	–
3	21.4	78.6	28.0	4.7	29.5	28.9	3.5	2.5	2.9	11.2	77.7	3.1
	—-	++++	–	–	++	++++	++++	—-	—-	++	++	++++
4	13.5	86.5	13.6	58.0	6.4	14.7	5.6	0.7	1.0	71.7	24.5	2.3
	—-	++++	—-	++++	—-	++	++++	—-	—-	++++	—-	++

CLASS	no facilities	electricity no	electricity yes	radio no	radio yes	tv no	tv yes	fridge no	fridge yes	bike no	bike yes	scooter no
1	50.4	98.1	1.9	63.0	37.0	98.4	1.6	99.8	0.2	74.1	25.9	92.9
	++	++	—-	++	–	++	—-		—-			
2	21.6	90.3	9.7	24.4	75.6	87.8	12.2	99.2	0.8	54.3	45.7	78.2
	–	++	—-	–	++	++	—-		—-	–	++	
3	8.0	16.9	83.1	23.3	76.7	19.1	80.9	77.5	22.5	80.3	19.7	72.6
	—-	—-	++++	–	++	—-	++++		++++		–	
4	1.5	4.2	95.8	17.8	82.2	5.7	94.3	24.9	75.1	75.6	24.4	84.3
	—-	—-	++++	—-	++	—-	++++	—-	++++		–	

CLASS	scooter yes	car no	car yes	floor earth	floor wood	floor cement	floor carpet	wall mud	wall cane	wall bricks	wall cement	wall metal
1	7.1	99.8	0.2	87.1	1.8	10.9	0.2	52.1	30.1	15.0	1.0	1.9
	—-		—-	++		—-	—-	++	++	—-	—-	–
2	21.8	98.4	1.6	48.6	1.3	49.4	0.6	35.7	9.5	39.0	13.5	2.4
	++		—-		–		—-		–	++		
3	27.4	95.0	5.0	9.6	1.5	81.3	7.6	7.9	11.4	37.6	39.9	3.2
	++			—-		++	++++	—-	–	++	++++	++
4	15.7	56.5	43.5	2.3	2.9	88.0	6.9	1.3	1.8	62.8	31.3	2.9
		–	++++	—-	++	++++	++++	—-	—-	++++	++++	

CLASS	roof thatch	roof tiles	roof metal	newsp never	newsp <onceaw	newsp onceaw	newsp everyday	radio never	radio<onceaw	radio onceaw	radio everyday	tv never
1	74.2	4.0	21.8	94.9	3.5	1.4	0.2	59.7	17.9	18.0	4.5	90.8
	++	—-	—-		—-	—-	—-	++		–	—-	++
2	24.8	9.5	65.6	84.3	10.4	4.6	0.7	25.4	20.4	33.5	20.7	73.1
	–		++			–	–	–			++	
3	3.1	12.2	84.7	71.8	16.2	10.3	1.7	24.0	20.5	41.1	14.3	15.2
	—-		++		++	++	++	–		++		—-
4	1.9	44.8	53.3	35.6	22.7	32.8	8.8	18.3	16.1	44.9	20.7	6.1
	—-	++++		—-	++++	++++	++++	—-		++	++	—-

CLASS	tv <onceaw	tv onceaw	tv everyday	poor	middle	rich	agricultural	manual	service	domestic	professional	unemployed
1	6.4	2.5	0.3	89.2	6.6	4.2	71.1	10.5	8.2	1.0	3.6	5.6
	–	—-	—-	++++	—-	—-	++	–	–	—-	—-	–
2	16.8	8.3	1.8	19.6	45.5	34.9	43.2	22.1	15.7	2.3	6.8	10.0
	++	—-	—-	—-	++++					–	–	++
3	16.6	46.7	21.6	4.9	9.4	85.6	8.7	34.5	29.3	5.7	14.6	7.1
	++	++++	++++	—-	—-	++++	—-	++	++	++	++	
4	9.4	55.0	29.6	0.2	2.0	97.9	3.9	26.0	22.3	8.2	34.2	5.3
	–	++++	++++	—-	—-	++++	—-	++	++	++++	++++	–

CLASS	agricultural	manual	service	domestic	professional	unemployed	Muslim	catholic	protestant	other Christian	African	no religion
1	54.7	5.0	12.7	0.5	0.5	26.6	30.1	17.8	18.6	21.2	3.8	8.4
	++	–	–	—-	—-						++	++
2	39.2	8.9	23.4	1.0	2.2	25.3	33.1	19.1	19.6	20.3	2.3	5.6
		++		–	–							
3	5.3	9.9	47.0	3.2	4.2	30.3	34.2	19.0	17.5	24.8	0.8	3.8
	—-	++	++	++	++						—-	–
4	2.1	6.9	29.8	8.2	21.4	31.5	23.5	20.0	29.0	21.0	0.2	6.3
	—-		++	++++	++++		–		++		—-	

Table 3.7: Area Classes Profiles

CLASS	rural	urban	well	piped-inside	public	piped-outside	bottled	spring	river,lake	toilet	pit-latrine	bucket
1	5.4	0.2	1.2	0.0	0.5	0.1	0.0	2.1	1.7	0.0	3.1	0.1
	++	—-	—-	—-	–	—-	—-	++++	++++	—-		
2	1.0	4.8	1.7	1.1	1.2	1.3	0.3	0.1	0.2	0.8	3.6	0.2
	—-	++++	–	++++		++++	++++	—-	—-	++++		++
3	1.0	4.9	0.9	0.8	2.3	1.3	0.2	0.2	0.1	0.5	4.0	0.3
	—-	++++	—-	++++	++++	++++	++	—-	—-	++++		++++
4	5.0	0.6	3.8	0.0	1.0	0.1	0.0	0.2	0.4	0.0	3.5	0.0
	++	—-	++	—-		—-	—-	—-	–	—-		—-

CLASS	no facility	electricity no	electricity yes	radio no	radio yes	tv no	tv yes	fridge no	fridge yes	bike no	bike yes	scooter no
1	2.4	5.4	0.2	3.1	2.5	5.4	0.3	5.6	0.0	4.7	0.9	5.3
	++	++	—-	++	–	++	—-		—-		–	
2	0.6	1.2	4.6	1.4	4.4	1.6	4.1	3.9	1.8	4.6	1.1	4.2
	—-	—-	++++	–	++	—-	++++	–	++++		–	
3	0.2	1.7	4.0	1.6	4.1	1.9	3.8	4.0	1.7	4.6	1.1	5.2
	—-	—-	++++	–		—-	++++		++++		–	
4	2.0	5.1	0.5	2.1	3.5	4.9	0.7	5.5	0.1	3.0	2.6	4.4
	++	++	—-			++	—-		—-	–	++	

CLASS	scooter yes	car no	car yes	floor earth	floor wood	floor cement	floor carpet	wall mud	wall cane	wall bricks	wall cement	wall metal
1	0.3	5.6	0.0	4.5	0.2	0.9	0.0	2.0	2.5	0.6	0.1	0.1
	—-		—-	++	++	—-	—-		++++	—-	—-	–
2	1.6	5.0	0.8	0.8	0.0	4.5	0.4	0.5	0.4	1.3	3.3	0.2
	++		++++	—-	—-	++	++++	—-	—-	–	++++	++
3	0.6	5.0	0.7	0.7	0.4	4.4	0.3	0.5	0.7	3.9	0.0	0.2
	–		++++	—-	++++	++	++	—-	–	++++	—-	++
4	1.2	5.5	0.1	3.6	0.0	1.9	0.0	2.7	0.4	1.8	0.5	0.1
	++		—-	++	—-		—-	++	—-		–	

CLASS	roof thatch	roof tiles	roof metal	newsp never	newsp <onceaw	newsp onceaw	newsp everyday	radio never	radio <onceaw	radio onceaw	radio everyday	tv never
1	3.4	0.5	1.6	5.1	0.3	0.1	0.0	3.0	1.0	1.3	0.3	4.8
	++		–		–	—-	—-	++		–	—-	++
2	0.3	1.2	4.2	4.1	0.9	0.8	0.0	1.4	1.4	3.1	0.0	1.2
	—-	++++	++		++	++++	—-	–	++	++	—-	—-
3	0.3	0.8	4.3	3.3	1.1	0.9	0.5	1.5	0.7	1.0	2.7	1.5
	—-	++	++	–	++++	++++	++++	–	–	–	++++	—-
4	2.6	0.3	2.6	5.0	0.4	0.2	0.0	2.1	1.1	1.5	0.9	4.4
	++	—-			–	–	—-					++

CLASS	tv <onceaw	tv onceaw	tv everyday	poor	middle	rich	on premises	<1hour	1to2hours	2to3hours	3+hours
1	0.5	0.2	0.1	3.9	1.0	0.7	1.4	3.3	0.3	0.1	0.1
	–	—-	—-	++		—-		++	++	++	++++
2	1.1	3.4	0.0	0.3	0.8	4.7	1.4	1.3	0.1	0.0	0.0
	++	++++	—-	—-	–	++++		–	—-	—-	—-
3	0.5	0.7	3.2	0.7	0.8	4.4	1.8	1.5	0.1	0.0	0.0
	–	–	++++	—-	–	++++		–	—-	—-	—-
4	0.7	0.5	0.1	3.0	1.4	1.3	1.7	2.8	0.2	0.0	0.0
		—-	—-	++	++	–					–

Appendix 2

Table 3.8: Cox Regression Results: Database with intervals up to 10 years before the survey.

Covariate	*coef*	*exp(coef)*	*se(coef)*
Parity			
1-2	0,000	1,000	*reference*
3-4	0,003	1,003	0,004
5-6	0,003	1,003	0,006
7-9	-0,036	0,965 ***	0,008
10+	-0,256	0,774 ***	0,016
Survived Children			
0-2	0,000	1,000	*reference*
3-5	-0,074	0,928 ***	0,005
6+	-0,139	0,870 ***	0,008
Previous Child is			
alive or dead since 1+ years	0,000	1,000	*reference*
dead $<$ 1 year	0,728	2,072 ***	0,005
Use of Birth Control			
No	0,000	1,000	*reference*
Yes	-2,700	0,067 ***	0,024
Mother's Age at Childbirth			
20-29	0,000	1,000	*reference*
$<$ 20	-0,015	0,985 ***	0,004
30-39	-0,252	0,777 ***	0,003
40+	-1,187	0,305 ***	0,010
Age at First Sex			
$<$ 15	0,000	1,000	*reference*
15-19	0,103	1,109 ***	0,005
20+	-0,058	0,943 ***	0,004
Difference with Husband Age			
0-4	0,000	1,000	*reference*
5-9	0,014	1,014 ***	0,003
10+	-0,024	0,977 ***	0,003
$<$ 0	-0,067	0,936 ***	0,006
unknown	-0,197	0,821 ***	0,005
Relation to Head			
Wife	0,000	1,000	*reference*
Head	-0,115	0,891 ***	0,003
Daughter	-0,244	0,783 ***	0,011
Other	-0,166	0,847 ***	0,006

Likelihood ratio test=161997 on 71 df

n= 1735978, number of events= 774661

Countries Dummies are omitted

Covariate	*coef*	*exp(coef)*	*se(coef)*
Cluster Analysis: Household Level			
Cluster 1	0,000	1,000	*reference*
Cluster 2	-0,034	0,966 ***	0,003
Cluster 3	-0,110	0,896 ***	0,005
Cluster 4	-0,176	0,839 ***	0,007
Female Education (years)			
0-2	0,000	1,000	*reference*
3-6	-0,033	0,968 ***	0,003
7-9	-0,093	0,911 ***	0,005
9+	-0,198	0,821 ***	0,005
Male Education (years)			
0-2	0,000	1,000	*reference*
3-6	-0,015	0,986 ***	0,003
7-9	-0,022	0,978 ***	0,005
9+	-0,071	0,932 ***	0,004
NA	-0,055	0,947 ***	0,007
Religion			
Muslim	0,000	1,000	*reference*
Catholic	-0,037	0,964 ***	0,004
Protestant	-0,021	0,979 ***	0,005
Other Catholic	-0,040	0,961 ***	0,004
Traditional	0,024	1,024 ***	0,008
Other	-0,017	0,984 ***	0,006
Cluster Analysis: Area Level			
Cluster 1	0,000	1,000	*reference*
Cluster 2	-0,001	0,999	0,006
Cluster 3	-0,067	0,935	0,007
Cluster 4	0,016	1,016 ***	0,003
Average Area Female Education	-0,013	0,987 ***	0,001
Stage of Demographic Transition			
Pre-Transition	0,000	1,000	*reference*
Early transition	-0,128	0,880 ***	0,004
Mid Transition	-0,221	0,802 ***	0,005
Late/Post-Transition	-0,341	0,711 ***	0,018

Likelihood ratio test=161997 on 71 df

n= 1735978, number of events= 774661

Countries Dummies are omitted

Table 3.9: Cox Regression Results: Database with intervals up to 3 years before the survey.

Covariate	*coef*	*exp(coef)*	*se(coef)*
Parity			
1-2	0,000	1,000	*reference*
3-4	-0,070	0,933 ***	0,016
5-6	-0,061	0,941 ***	0,020
7-9	-0,069	0,934 ***	0,026
10+	-0,169	0,844 ***	0,045
Survived Children			
0-2	0,000	1,000	*reference*
3-5	-0,096	0,908 ***	0,016
6+	-0,120	0,887 ***	0,024
Previous Child is			
alive or dead since 1+ years	0,000	1,000	*reference*
dead < 1 year	0,954	2,597 ***	0,017
Use of Birth Control			
No	0,000	1,000	*reference*
Yes	-3,560	0,028 ***	0,055
Mother's Age at Childbirth			
20-29	0,000	1,000	*reference*
< 20	0,014	1,014	0,015
30-39	-0,302	0,740 ***	0,011
40+	-1,199	0,301 ***	0,023
Age at First Sex			
< 15	0,000	1,000	*reference*
15-19	0,117	1,124 ***	0,016
20+	-0,028	0,972 *	0,016
Difference with Husband Age			
0-4	0,000	1,000	*reference*
5-9	-0,021	0,979 **	0,010
10+	-0,065	0,937 ***	0,011
< 0	0,011	1,011	0,021
unknown	-0,616	0,540 ***	0,020
Relation to Head			
Wife	0,000	1,000	*reference*
Head	-0,142	0,868 ***	0,012
Daughter	-0,111	0,895 ***	0,028
Other	-0,217	0,805 ***	0,020
Likelihood ratio test=35274 on 71 df			
n= 317519, number of events= 65913			
Countries Dummies are omitted			

Covariate	*coef*	*exp(coef)*	*se(coef)*
Cluster Analysis: Household Level			
Cluster 1	0,000	1,000	*reference*
Cluster 2	-0,057	0,945 ***	0,010
Cluster 3	-0,118	0,889 ***	0,018
Cluster 4	-0,148	0,862 ***	0,027
Female Education (years)			
0-2	0,000	1,000	*reference*
3-6	0,006	1,006	0,011
7-9	-0,049	0,952 **	0,017
9+	-0,110	0,896 ***	0,018
Male Education (years)			
0-2	0,000	1,000	*reference*
3-6	0,020	1,020 *	0,012
7-9	-0,007	0,993	0,016
9+	-0,054	0,948 ***	0,014
NA	-0,018	0,982	0,024
Religion			
Muslim	0,000	1,000	*reference*
Catholic	-0,026	0,975 *	0,015
Protestant	0,007	1,007	0,015
Other Catholic	-0,015	0,985	0,014
Traditional	-0,019	0,981	0,029
Other	0,027	1,027	0,020
Cluster Analysis: Area Level			
Cluster 1	0,000	1,000	*reference*
Cluster 2	0,020	1,020	0,022
Cluster 3	-0,119	0,888 ***	0,025
Cluster 4	0,004	1,004	0,012
Average Area Female Education	0,004	1,004	0,003
Stage of Demographic Transition			
Pre-Transition	0,000	1,000	*reference*
Early transition	0,034	1,035	0,025
Mid Transition	-0,060	0,941 ***	0,029
Late/Post-Transition	-0,231	0,794 ***	0,048

Likelihood ratio test=35274 on 71 df

n= 317519, number of events= 65913

Countries Dummies are omitted

Table 3.10: Cox Regression Results: Database with intervals up to 5 years before the survey.

Covariate	*coef*	*exp(coef)*	*se(coef)*
Parity			
1-2	0,000	1,000	*reference*
3-4	-0,070	0,933 ***	0,015
5-6	-0,034	0,967 *	0,018
7-9	-0,025	0,975	0,024
10+	-0,149	0,862 ***	0,039
Survived Children			
0-2	0,000	1,000	*reference*
3-5	-0,102	0,903 ***	0,015
6+	-0,126	0,882 ***	0,022
Previous Child is			
alive or dead since 1+ years	0,000	1,000	*reference*
dead < 1 year	1,097	2,995 ***	0,015
Use of Birth Control			
No	0,000	1,000	*reference*
Yes	-3,272	0,038 ***	0,054
Mother's Age at Childbirth			
20-29	0,000	1,000	*reference*
< 20	0,028	1,028 *	0,012
30-39	-0,291	0,748 ***	0,011
40+	-1,104	0,332 ***	0,024
Age at First Sex			
< 15	0,000	1,000	*reference*
15-19	0,102	1,108 ***	0,015
20+	-0,050	0,952 ***	0,014
Difference with Husband Age			
0-4	0,000	1,000	*reference*
5-9	-0,033	0,968 ***	0,009
10+	-0,056	0,945 ***	0,010
< 0	0,015	1,015	0,018
unknown	-0,505	0,604 ***	0,018
Relation to Head			
Wife	0,000	1,000	*reference*
Head	-0,138	0,871 ***	0,010
Daughter	-0,138	0,871 ***	0,027
Other	-0,183	0,832 ***	0,018
Likelihood ratio test=33037 on 71 df			
n= 327527, number of events= 83109			
Countries Dummies are omitted			

Covariate	*coef*	*exp(coef)*	*se(coef)*
Cluster Analysis: Household Level			
Cluster 1	0,000	1,000	*reference*
Cluster 2	-0,057	0,945 ***	0,009
Cluster 3	-0,118	0,888 ***	0,016
Cluster 4	-0,198	0,820 ***	0,028
Female Education (years)			
0-2	0,000	1,000	*reference*
3-6	0,000	1,000	0,010
7-9	-0,055	0,947 ***	0,015
9+	-0,116	0,890 ***	0,018
Male Education (years)			
0-2	0,000	1,000	*reference*
3-6	0,012	1,012	0,010
7-9	-0,027	0,973 **	0,014
9+	-0,062	0,940 ***	0,012
NA	-0,020	0,980	0,021
Religion			
Muslim	0,000	1,000	*reference*
Catholic	-0,043	0,958 ***	0,013
Protestant	-0,017	0,984	0,014
Other Catholic	-0,027	0,974 *	0,013
Traditional	0,021	1,021	0,024
Other	0,007	1,007	0,017
Cluster Analysis: Area Level			
Cluster 1	0,000	1,000	*reference*
Cluster 2	0,020	1,021	0,020
Cluster 3	-0,114	0,892 ***	0,023
Cluster 4	0,027	1,028 **	0,010
Average Area Female Education	-0,003	0,997	0,003
Stage of Demographic Transition			
Pre-Transition	0,000	1,000	*reference*
Early transition	-0,129	0,879 ***	0,021
Mid Transition	-0,212	0,809 ***	0,025
Late/Post-Transition	-0,383	0,682 ***	0,052

Likelihood ratio test=33037 on 71 df

n= 327527, number of events= 83109

Countries Dummies are omitted

Appendix 3

To test the proportionality assumption we used the command "cox.zph" in R. A p-value lower than 5% indicates that the coefficient are time dependent and therefore the assumption is violated. Therefore higher is the p-value, higher is the probability that the assumption is respected. In the best case scenario the p-value is higher than 5% not only for the single categories of the single variables, but also for the model, globally. In our case we cannot reject the null hypothesis globally, however the categories of the main variables have a p-value close to, or higher than, 5%, leading to the conclusion that there is not a strong violation of the proportionality assumption.

Table 3.11: Proportionality test

Variable	rho	chisq	p-value
Parity			
3-4	0,010	8,217	0,004
5-6	0,009	6,507	0,011
7-9	-0,005	2,131	0,144
10+	-0,018	29,452	0,000
Survived Children			
3-5	-0,002	0,299	0,584
6+	-0,005	2,112	0,146
Last Child is dead	-0,150	1937,068	0,000
Birth Control Yes	-0,009	6,650	0,010
Mother Age			
< 20	0,004	1,494	0,222
30-39	-0,003	0,783	0,376
40+	-0,019	33,161	0,000
Age at First Sex			
15-19	-0,016	23,961	0,000
20+	-0,013	15,324	0,000
Difference with Husband Age			
5-9	0,002	0,242	0,623
10+	-0,002	0,363	0,547
< 0	-0,009	7,736	0,005
NA	-0,033	96,456	0,000
Relation to Head			
Head	-0,009	0,007	0,005
Daughter	-0,002	0,478	0,491
Other	-0,004	0,197	0,164

Variable	rho	chisq	p-value
Parity			
Cluster Analysis: Household Level			
Cluster 2	-0,005	0,109	0,121
Cluster 3	-0,012	0,000	0,000
Cluster 4	-0,015	0,000	0,000
Female Education (Years)			
3-6	0,009	0,007	0,007
7-9	0,003	0,419	0,352
9+	-0,001	0,819	0,857
Male Education (Years)			
3-6	0,001	0,136	0,712
7-9	0,003	0,565	0,452
9+	0,003	0,787	0,375
NA	-0,003	0,727	0,394
Religion			
Catholic	0,008	6,352	0,012
Protestant	0,003	0,649	0,420
Other Christian	0,003	0,910	0,340
Traditional	0,001	0,125	0,724
Other	-0,001	0,118	0,731
Cluster Analysis: Area Level			
Cluster 2	-0,007	3,870	0,049
Cluster 3	-0,005	2,453	0,117
Cluster 4	-0,003	0,779	0,378
Average Female Edu	-0,017	25,913	0,000
Stage of Demographic Transition			
Early Transition	0,019	37,023	0,000
Mid Transition	0,015	21,554	0,000
Late/Post Transition	0,010	8,486	0,004
GLOBAL	NA	4064,557	0,000

Appendix 4

The variance inflation factor (VIF) is a measure of correlation between the main independent variables. It is obtained by regressing one of the explanatory variables over the others. We calculated the VIF for the variables in our model using the command "vif" in R. The fact that we are using a linear model in a framework of survival analysis must not be worrisome. The VIF refers to the correlation between the explicative variables only, therefore can be applied in many different circumstances, no matter the final method of analysis.

Table 3.12: VIF: Variance Inflation Factor

Variables	Cathegories				
Parity	3-4	5-6	7-9	10+	
VIF	4,60	4,85	5,00	2,00	
Survived Children	3-5	6+			
VIF	5,13	5,05			
Last Child is	dead				
VIF	1,04				
Birth Control Yes	Yes				
VIF	1,08				
Mother Age	< 20	30-39	40+		
VIF	1,32	1,75	1,60		
Age at First Sex	15-19	20+			
VIF	1,21	1,03			
Difference with Husband Age	5-9	10+	< 0	NA	
VIF	1,46	1,59	1,13	1,50	
Relation to Head	Head	Daughter	Other		
VIF	1,32	1,19	1,18		
Cluster Analysis: Household Level	Cluster 2	Cluster 3	Cluster 4		
VIF	1,46	2,91	2,00		
Female Education (Years)	3-6	7-9	9+		
VIF	1,55	1,54	2,20		
Male Education (Years)	3-6	7-9	9+	NA	
VIF	1,56	1,51	2,39	1,14	
Cluster Analysis: Area Level	Cluster 2	Cluster 3	Cluster 4		
VIF	3,78	2,96	2,36		
Stage of Demographic Transition	Early Transition	Mid Transition	Late/Post Transition		
VIF	5,99	7,17	2,23		
Religion	Catholic	Protestant	Other Christian	Traditional	Other
VIF	2,19	2,51	2,10	1,20	1,53
Country	Burkina Faso	Burundi	Cameroon	Chad	Comoros
VIF	2,20	1,71	2,07	1,53	1,33
Country	Congo	Congo DR	Ethiopia	Gabon	Ghana
VIF	1,98	3,37	2,41	1,62	1,27
Country	Guinea	Cote d'Ivoire	Kenya	Lesotho	Liberia
VIF	1,58	1,55	1,66	1,59	1,67
Country	Madagascar	Malawi	Mali	Mozambique	Namibia
VIF	2,26	2,97	2,17	1,89	1,59
Country	Nigeria	Rwanda	Senegal	Sierra Leone	Swaziland
VIF	3,43	2,33	2,14	1,98	1,28
Country	Togo	Uganda	Zambia	Zimbawe	
VIF	1,58	1,85	2,93	1,84	
Global VIF	2,22				

Appendix 5

Table 3.13: Cox Regression Results: Effect of the levels of education and of the other covariates on the risk of second and higher order pregnancies (Hazard Ratio), for women aged 15-49, in the 5 years before the survey. Regression with no Stage of Demographic Transition

Covariate	*coef*	*exp(coef)*	*se(coef)*
Parity			
1-2	0,000	1,000	*reference*
3-4	-0,086	0,917 ***	0,015
5-6	-0,064	0,938 ***	0,019
7-9	-0,068	0,934 **	0,024
10+	-0,220	0,803 ***	0,040
Survived Children			
0-2	0,000	1,000	*reference*
3-5	-0,100	0,904 ***	0,015
6+	-0,136	0,873 ***	0,022
Previous Child is			
alive or dead since 1+ years	0,000	1,000	*reference*
dead < 1 year	1,111	3,038 ***	0,016
Use of Birth Control			
No	0,000	1,000	*reference*
Yes	-3,283	0,038 ***	0,054
Mother's Age at Childbirth			
20-29	0,000	1,000	*reference*
< 20	0,022	1,023 *	0,012
30-39	-0,282	0,755 ***	0,011
40+	-1,096	0,334 ***	0,024
Age at First Sex			
< 15	0,000	1,000	*reference*
15-19	0,103	1,109 ***	0,015
20+	-0,049	0,953 **	0,015
Difference with Husband Age			
0-4	0,000	1,000	*reference*
5-9	-0,032	0,968 **	0,009
10+	-0,058	0,944 ***	0,010
< 0	0,013	1,013	0,019
unknown	-0,512	0,599 ***	0,018

Events, n = 83109, 327527

Integrated Log-likelihood = -924776

Integrated Log-likelihood Chisq = 33905.28 df 69

Frailty Model (RE) at Area Level

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

Covariate	*coef*	*exp(coef)*	*se(coef)*
Relation to Head			
Wife	0,000	1,000	*reference*
Head	-0,147	0,863 ***	0,011
Daughter	-0,134	0,875 ***	0,028
Other	-0,194	0,824 ***	0,018
Cluster Analysis: Household Level			
Cluster 1	0,000	1,000	*reference*
Cluster 2	-0,060	0,941 ***	0,009
Cluster 3	-0,136	0,873 ***	0,017
Cluster 4	-0,217	0,805 ***	0,029
Female Education (years)			
0-2	0,000	1,000	*reference*
3-6	-0,001	0,999	0,010
7-9	-0,054	0,947 ***	0,015
9+	-0,124	0,883 ***	0,019
Male Education (years)			
0-2	0,000	1,000	*reference*
3-6	0,013	1,013	0,011
7-9	-0,026	0,974 *	0,014
9+	-0,069	0,933 ***	0,013
NA	-0,022	0,978	0,022
Religion			
Muslim	0,000	1,000	*reference*
Catholic	-0,033	0,968 *	0,015
Protestant	-0,005	0,995	0,016
Other Catholic	-0,023	0,978	0,015
Traditional	0,023	1,023	0,026
Other	0,022	1,022	0,019
Cluster Analysis: Area Level			
Cluster 1	0,000	1,000	*reference*
Cluster 2	-0,046	0,955 *	0,022
Cluster 3	-0,184	0,832 ***	0,024
Cluster 4	0,011	1,011	0,013
Average Area Female Education	-0,007	0,993 *	0,003

Events, n = 83109, 327527

Integrated Log-likelihood = -924776

Integrated Log-likelihood Chisq = 33905.28 df 69

Frailty Model (RE) at Area Level

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

Covariate	*coef*	*exp(coef)*	*se(coef)*
Country			
Benin	0	1	reference
Burkina Faso	-0,048	0,954 *	0,028
Burundi	0,468	1,597 ***	0,033
Cameroon	0,468	1,596 ***	0,031
Chad	0,445	1,561 ***	0,037
Comoros	0,552	1,738 ***	0,044
Congo	0,433	1,542 ***	0,036
Congo DR	0,636	1,888 ***	0,030
Ethiopia	0,179	1,196 ***	0,030
Gabon	0,680	1,974 ***	0,040
Ghana	0,100	1,105 *	0,046
Guinea	-0,115	0,892 ***	0,035
Cote d'Ivoire	0,092	1,097 **	0,034
Kenya	0,549	1,731 ***	0,039
Lesotho	-0,122	0,885 **	0,049
Liberia	0,119	1,127 ***	0,035
Madagascar	0,262	1,300 ***	0,031
Malawi	0,227	1,255 ***	0,028
Mali	0,120	1,127 ***	0,030
Mozambique	0,213	1,238 ***	0,030
Namibia	0,174	1,190 ***	0,054
Nigeria	0,364	1,439 ***	0,026
Rwanda	0,227	1,255 ***	0,034
Senegal	0,275	1,316 ***	0,031
Sierra Leone	0,025	1,025	0,031
Swaziland	0,404	1,498 ***	0,060
Togo	-0,015	0,985	0,035
Uganda	0,640	1,897 ***	0,033
Zambia	0,372	1,451 ***	0,032
Zimbabwe	-0,016	0,984	0,050

Events, n = 83109, 327527

Integrated Log-likelihood = -924776

Integrated Log-likelihood Chisq = 33905.28 df 69

Frailty Model (RE) at Area Level

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

Bibliography

[1] Akmam, W. (2002). Women's education and fertility rates in developing countries, with special reference to Bangladesh. *Eubios Journal of Asian and International Bioethics*, 12.4: 138-143.

[2] Alter, G. (2016). Understanding Historical and Contemporary Fertility Transitions: A Birth Interval Approach. *Unpublished IUSSP Working Paper (quoted with author's permission).*

[3] Anderton, D.L. et Bean, L.L. (1985). Birth spacing and fertility limitation: A behavioral analysis of a nineteenth-century frontier population. *Demography*, 22, pp. 169-183.

[4] Angeles, G., Guilkey, D.K. and Mroz, T.A. (2005). The effects of education and family planning programs on fertility in Indonesia. *Economic Development and Cultural Change*, 54.1 (2005): 165-201.

[5] Appadurai, A.(2004). The capacity to aspire. *Culture and public action* (2004): 59-84.

[6] Aurig, G. (2013). Persistent High Fertility in Sub Saharan Africa. *LUP Student Papers,* 3866778.

[7] Bankole, A., Audam, S. (2011). Fertility preferences and contraceptive use among couples in sub-Saharan Africa. *African Population Studies*, 25(2).

[8] Basu, A.M. (2002). Why does education lead to lower fertility? A critical review of some of the possibilities. *World Development* 30(10): 1779-1790.

[9] Becker, A., Deckers, T., Dohmen, T., Falk, A., Kosse, F. (2012). The relationship between economic preferences and psychological personality measures. *Annu. Rev. Econ.*, 4(1), 453-478.

[10] Becker, G. S. (1968). Crime and punishment: An economic approach. *The Economic Dimensions of Crime (pp. 13-68). Palgrave Macmillan UK.*

Becker, G. S. (1993). Nobel lecture: The economic way of looking at behavior. *Journal of political economy*, 101(3), 385-409.

Becker, G. S. (2013). The economic approach to human behavior. *University of Chicago press.*

[11] Behrman, J. A. (2015). Does schooling affect women's desired fertility? Evidence from Malawi, Uganda, and Ethiopia. *Demography*, 52(3), 787-809.

[12] Benabou, R. (2002). Tax and education policy in a heterogeneous' agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica, 70(2), 481-517.*

[13] Benefo, K. D., Schultz, T. P. (1994). Determinants of fertility and child mortality in Cote d'Ivoire and Ghana (Vol. 103). *World Bank Publications.*

[14] Bledsoe, et al. (1998). Critical perspectives on schooling and fertility in the developing world. *National Academies Press.*

[15] Bongaarts, J. (2003). Completing the fertility transition in the developing world: The role of educational differences and fertility preferences. *Population Studies*, 57(3), 321-335.

Bongaarts, J. (2010). The causes of educational differences in fertility in Sub-Saharan Africa. *Vienna yearbook of population research* ,31-50.

Bongaarts, J. (2017). Africa's unique fertility transition. *Population and Development Review*, 43(S1), 39-58.

[16] Bongaarts, J., Casterline, J. (2013). Fertility transition: is sub Saharan Africa different? *Population and development review*, 38(s1), 153-168.

[17] Bourguignon, F. (1999). Crime, violence and inequitable development. *Annual World Bank Conference on Development Economics* 1999 (pp. 199-220).

[18] Bowles, S. (1998). Endogenous preferences: The cultural consequences of markets and other economic institutions. *Journal of economic literature*, 36(1), 75-111.

[19] Brass, W., Jolly, C.L. (1993). Population Dynamics of Kenia. *Population Dynamics of Sub-Saharan Africa: Working Group on Kenia.*

[20] Bulatao, R.A, Lee, R.D. (1983). An Overview of Fertility Determinants in Developing Countries. Determinants of Fertility in Developing Countries, vol. 2: Fertility Regulation and Institutional Influences. *New York: Academic Press*, 757-787.

[21] Buonanno, P., Leonida, L. (2006). Education and crime: evidence from Italian regions. *Applied Economics Letters*, 13(11), 709-713.

Buonanno, P., Leonida, L. (2009). Non-market effects of education on crime: Evidence from Italian regions. *Economics of Education Review* 28.1 (2009): 11-17.

[22] Caldwell, J.C. (1982). Theory of Fertility Decline. *London: Academic Press.*

[23] Caldwell, J. C., Caldwell, P. (1987). The cultural context of high fertility in sub-Saharan Africa. *Population and development review*, 409-437.

[24] Caldwell, J.C., Orubuloye, I.O. and Caldwell, P. (1992). Fertility Decline in Africa: A New Type of Transition? *Population and Development Review*, 18(2):211-242.

[25] Canning, D., Raja, S., Yazbeck, A. S. (2015). Africa's demographic transition: dividend or disaster? *World Bank Publications.*

[26] Carter, A.T. (1999) . What Is Meant, and Measured, by Education? *Critical Perspectives on Schooling and Fertility in the Developing World. Washington, DC: National Academy Press*, 49-79.

[27] Casterline, J.B., and El Zeini, L.O. (2014). Unmet Need and Fertility Decline: A Comparative Perspective on Prospects in Sub-Saharan Africa. *Studies in family planning* 45.2 (2014): 227-245.

[28] Casterline, J.B., Williams, L. and McDonald, P. (2010). The Age Difference Between Spouses: Variations among Developing Countries. *Population Studies: A Journal of Demography*, 40:3, 353-374.

[29] Casterline, J.B. (2001). Diffusion processes and fertility transition: selected perspectives *National Research Council*.

Casterline, J.B. (2010). Determinants and Consequences of High Fertility: A Synopsis of the Evidence. *World Bank Report available on the website: http://www.worldbank.org/hnppublications*.

[30] Castro Martín, T. (1995). Womens, Education and Fertility: Results from 26 Demographic and Health Surveys. *Studies in Family Planning* 26(4): 187-202.

[31] Castro Martín, T., Juárez, F. (1995). The impact of women's education on fertility in Latin America: Searching for explanations. *International Family Planning Perspectives*. 21(2):52-57.

[32] Cinnirella, F., Klemp, M., and Weisdorf, J. (2012). Malthus in the Bedroom: Birth Spacing as a Preventive Check Mechanism in Pre-Modern England. *CEPR Discussion Paper No. DP9116*.

Cinnirella, F., Klemp, M., and Weisdorf, J. (2017). Malthus in the bedroom: birth spacing as birth control in pre-transition England. *Demography*, 54(2), 413-436.

[33] Cochrane, S.H. (1979). Fertility and Education: What Do We Really Know? *Baltimore, Md.: Johns Hopkins University Press*.

[34] Cohen, B. (1998). The emerging fertility transition in Sub-Saharan Africa. *World Development*, 26(8), 1431-1461.

[35] Conde-Agudelo, A., Rosas-Bermudez, A., Castaño, F. and Norton, M.H. (2012). Effects of Birth Spacing on Maternal, Perinatal, Infant, and Child Health: A Systematic Review of Causal Mechanisms. *Studies in Family Planning.* 43(2):93-114.

[36] Dasgupta, P. (2004). World poverty: Causes and pathways. *In Annual World Bank Conference on Development Economics* (pp. 159-223).

[37] De Baun, D., Roc, M.(2013). Saving futures, savings dollars: The impact of education on crime reduction and earnings. *Retrieved from Alliance for Excellent Education*: http://all4ed.org/wpcontent/uploads/2013/09/SavingFutures.pdf

[38] Decker, S. H., Ortiz, N., Spohn, C., and Hedberg, E. (2015). Criminal stigma, race, and ethnicity: The consequences of imprisonment for employment. *Journal of Criminal Justice*, 43(2), 108-121.

[39] De La Croix, D., Doepke, M. (2003). Inequality and growth: why differential fertility matters. *American Economic Review*, 93(4), 1091-1113.

[40] De la Croix, D., Gobbi, P. (2014). L'expansion de l'éducation en Afrique annonce-t-elle sa transition démographique? *Dounia*. 7:33-50.

De la Croix, D., Gobbi, P. (2017). Population density, fertility, and demographic convergence in developing countries. *Journal of Development Economics* 127 (2017): 13-24.

[41] DeRose, L.F., Kravdal, Ø.(2007). Educational reversals and first-birth timing in sub-Saharan Africa: A dynamic multilevel approach. *Demography* 44.1 (2007): 59-77.

[42] Detotto, C., Otranto, E. (2010). Does crime affect economic growth? *Kyklos*, 63(3), 330-345.

[43] Diamond, P.A. (1982). Aggregate demand management in search equilibrium. *Journal of Political Economy* 90 (5), 881-894.

Diamond, P.A. (1982). Wage determination and efficiency in search equilibrium. *Review of Economic Studies* 49 (2), 217-227.

[44] Diamond, I., Newby, M., and Varle, S. (1999). Female Education and Fertility: Examining the Links. *Critical Perspectives on Schooling and Fertility in the Developing World. Washington, DC: National Academy Press*, 287-306: 23-48.

[45] Drèze, J., Murthi, M. (2001). Fertility, education, and development: evidence from India. *Population and development Review*, 27(1), 33-63.

[46] Dyson, T., Murphy, M. (1985). The Onset of Fertility Transition. *Population and Development Review*, 11:399-440.

[47] Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *The Journal of Political Economy*, 521-565.

Ehrlich, I. (1975). On the relation between education and crime. *Education, income, and human behaviour* (pp. 313-338). NBER.

[48] Eloundou-Enyegue, P.M. (1999). Fertility And Education: What Do We Now Know? *In Bledsoe, C.H., Casterline J.B., Johnson-Kuhn J.A., and Haaga J.G. (eds.), Critical Perspectives on Schooling and Fertility in the Developing World. Washington, DC: National Academy Press*, 287-306.

[49] Fajnzylber, P., Lederman, D. and Loayza, N. (2002). Inequality and violent crime. *Journal of Law and Economics* 45 (1), 1-40.

[50] Fella, G., Gallipoli, G. (2008). Education and crime over the life cycle. *No. 630. Working Paper//School of Economics and Finance, Queen Mary, University of London,* 2008.

[51] Fenn, N.S., Edmeades, J., Lantos, H., and Onovo, O. (2915). Child marriage, Adolescent pregnancy and Family formation in West and Central Africa, *UNICEF*.

[52] Fernandez, R., Rogerson, R. (1998). Public education and income distribution: A dynamic quantitative evaluation of education-finance reform. *American Economic Review*, 813-833.

Fernandez, R., Rogerson, R. (1999). Equity and resources: An analysis of education finance systems. *National Bureau of Economic Research*, (No. w7111).

[53] Fisher, K. (2000). Uncertain aims and tacit negotiation: birth control practices in Britain, 1925-1950. *Population and Development Review*, 26, pp. 295-31

[54] Freedman, R., Chang, M.C. and Sun, T.H. (1994). Taiwan's transition from high fertility to below-replacement levels. *Studies in Family Planning* 25(6):317-331.

[55] Friedlander, D., Okun, B.S. and Segal, S. (1999). The demographic transition then and now: Processes, perspectives, and analyses. *Journal of Family History*, 24, pp. 493-53.

[56] Garenne, M. (2012). Education and Fertility in Sub-Saharan Africa: A Longitudinal Perspective. *DHS Analytical Studies No. 33. Calverton, Maryland, USA: ICF International.*

Garenne, M. (2014). Trends in Marriage and Contraception in Sub-Saharan Africa: A Longitudinal Perspective on Factors of Fertility Decline. *DHS Analytical Studies No. 42. Rockville, Maryland, USA: ICF International.*

[57] Gebreselassie, T. (2011). The Fertility Transition in Sub-Saharan Africa, 1990?2005: How Unique Is Ethiopia? *A chapter in Butz, William. The Demographic Transition and Development in Africa. Ed. Charles Teller. Springer Science+ Business Media BV.*

[58] Glomm, G., Ravikumar, B. (1992). Public versus private investment in human capital: endogenous growth and income inequality. *Journal of political economy* 100.4: 818-834.

[59] Godin, M., and Hindriks, J. (2016). Equité et efficacité des écoles: une comparaison internationale basée sur la mobilité sociale à l'école. *UCL-Université Catholique de Louvain.*

[60] Goulas, E., Zervoyianni, A. (2013). Economic growth and crime: does uncertainty matter? *Applied Economics Letters*, 20(5), 420-427.

[61] Harlow, C. W.(2003). Education and Correctional Population. *Special Reports, U.S. Department of Justice.*

[62] Hionidou, V. (1998). The adoption of fertility control on Mykonos, 1879-1959: stopping, spacing or both? *Population Studies*, 52, pp. 67-83.

[63] Hosmer, D.W., Lemeshow, S. (1999). Applied Survival Analysis: Regression Modeling of Time to Event Data. *New York: Wiley, Inc.*

[64] Jejeebhoy, S.J. (1995). Women's Education, Autonomy and Reproductive Behavior: Experience from Developing Countries. *Oxford: Clarendon Press.*

[65] Kaira, K., Takahashi, T., Murakami, H., Tsuya, A., Nakamura, Y., Naito, T., and Yamamoto, N. (2010). Long-term survivors of more than 5 years in advanced non-small cell lung cancer. *Lung Cancer*, 67(1), 120-123.

[66] Kenworthy, L., Pontusson, J. (2005). Rising inequality and the politics of redistribution in affluent countries. *Perspectives on Politics*, 3(3), 449-471.

[67] Knodel, J. (1987). Starting, stopping, and spacing during the early stages of fertility transition: The experience of German village populations in the 18th and 19th centuries. *Demography* 24(2):143-162.

[68] Knodel, J. and van de Walle, E. (1979). Lessons from the Past: Policy Implications of Historical fertility Studies. *Population and Development Review.* 5(2):217-245;

[69] Kravdal, Ø. (2002). Education and fertility in sub-Saharan Africa: Individual and community effects. *Demography.* 39: 233-250.

Kravdal, Ø. (2012). Further evidence of community education effects on fertility in sub-Saharan Africa. *Demographic Research,* 27: 645-680.

[70] Kumar, S. (2013). Crime and economic growth: evidence from India. *MPRA, Paper N.48794.*

[71] Lee, D. (2005). An Estimable Dynamic General Equilibrium Model Of Work, Schooling, And Occupational Choice. *International Economic Review, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association*, vol. 46(1), pages 1-34.

[72] Lesthaeghe, R., Vanderhoeft, C., Becker, and S., Kibet, M. (1985). Individual and contextual effects of education on proximate fertility determinants and on life-time fertility in Kenya. *In John B. Casterline (ed.).The Collection and Analysis of Community Data. Voorburg, The Netherlands: International Statistical Institute*, 31-63.

[73] Lesthaeghe, R., Permanyer, I. (2014). European sub-replacement fertility: trapped or recovering? *Population Studies Center Research Report*, (14-822).

[74] Lloyd, C.B. (1994). Investing in the Next Generation: The Implications of High Fertility at the Level of the Family. *Working Paper No. 63. New York: The Population Council.*

[75] Lloyd, C. B., Kaufman, C.E., and Hewett, P. (2000). The spread of primary schooling in sub-Saharan Africa: Implications for fertility change. *Population and Development Review* 26.3 (2000): 483-515.

[76] Lochner, L. (2004). Education, work, and crime: A human capital approach. *International Economic Review* 45.3 (2004): 811-843.

Lochner, L. (2007). Education and crime. *University of Western Ontario* 5.8 (2007): 1-14.

[77] Lochner, L., Moretti, E. (2003). The Effect of Education on Crime: Evidence from Prison Inmates, Arrests and Self-Reports. *American Economic Review, 94, 155-189.*

[78] Machin, S., Meghir, C. (2004). Crime and economic incentives. *Journal of Human Resources*, 39(4), 958-979.

[79] Machin, S., Marie, O., and Vujic, S. (2011). The Crime Reducing Effect of Education. *The Economic Journal* 121.552: 463-484.

[80] Mason, K.O. (1984). The Status of Women: A Review of Its Relationship to Fertility and Mortality. *NewYork: The Rockefeller Foundation.*

[81] Mauro, L., Carmeci, G. (2007). A poverty trap of crime and unemployment. *Review of Development Economics*, 11(3), 450-462.

[82] McDevitt, T., Johnson, P. (2005). Measurement issues and proximate Determinants of slow and stagnating fertility decline: Case studies of Kenya and the Philippines. *Extended Abstract. http://paa2005.princeton.edu/download.aspx?submissionId=50142* (Accessed 21 May 2018).

[83] McDonald, P. (1984). Nuptiality and Completed Fertility: A Study of Starting, Stopping, and Spacing Behavior. *Voorburg: International Statistical Institute.*

[84] Mocan, H. N., Billups, S. C., and Overland, J. (2005). A dynamic model of differential human capital and criminal activity. *Economica*, 72(288), 655-681.

[85] Monstad, K., Propper, C. and Salvanes, K. G. (2008). Education and Fertility: Evidence from a Natural Experiment. *The Scandinavian Journal of Economics*, 110: 827?852.

[86] Mortensen, D.T. (1982a). The matching process as a non-cooperative bargaining game. *McCall, J.J. (Ed.), The Economics of Information and Uncertainty. University of Chicago Press, New York.*

Mortensen, D.T. (1982b). Property rights and efficiency in mating, racing and related games. *American Economic Review* 72, 968-979.

[87] Mortensen, D.T., Pissarides, C.A. (1994). Job creation and job destruction in the theory of unemployment. *Review of Economic Studies* 61, 397?415.

[88] Moultrie, T.A., Sayi, T.S. and Timæus, I.M. (2012). Birth intervals, postponement, and fertility decline in Africa: A new type of transition? *Population Studies.* 66(3):241-258.

[89] Myrskylä, M. and Barclay, K. (2017). Fertility postponement could reduce child mortality: evidence from 228 demographic and health surveys covering 77 developing countries. *MPIDR Working Paper WP* 2017-005.

[90] Okun, B.S. (1995). Distinguishing stopping behavior from spacing behavior with indirect Methods. *Historical Methods*, 28, pp. 85-96.

[91] Park, S., Hendry, D. J. (2015). Reassessing Schoenfeld residual tests of proportional hazards in political science event history analyses. *American Journal of Political Science*, 59(4), 1072-1087.

[92] Persson, T., Tabellini, G. (1994). Is inequality harmful for growth? *The American economic review*, 600-621.

[93] Pissarides, C.A. (1979). Job matchings with state employment agencies and random search. *Economic Journal* 89, 818-833.

Pissarides, C.A. (1985). Short-run equilibrium dynamics of unemployment, vacancies and real wages. *American Economic Review* 75, 676-690.

[94] Powell, B., Manish, G. P., and Nair, M. (2010). Corruption, crime and economic growth. *Handbook on the Economics of Crime*, 328.

[95] Ray, D.(2003). Aspiration, Poverty and Economic Change. *mimeo, New York University.*

[96] RTE. RTE act. *available at www.mhrd.gov.in/rte.* Last access 27/05/2018.

[97] Rukmini, S. India officially undercounts all crimes including rape. *The Hindu*, available at http://www.thehindu.com/news/national/india-officially-undercounts-all-crimes-including-rape/article5121114.ece, last access 15/09/2017.

[98] Saint-Paul, G., Verdier, T. (1993). Education, democracy and growth. *Journal of Development Economics*, 42(2), 399-407.

Saint-Paul, G., Verdier, T. (1996). Inequality, redistribution and growth: A challenge to the conventional political economy approach. *European Economic Review*, 3(40), 719-728.

[99] Samari, G. (2017). Demographic Effects of Girls' Education in Developing Countries: Proceedings of a Workshop. In Brief. *National Academies Press*.

[100] Sandhu, J. (1996). Sociology of Fertility. *Jaipur and New Delhi: Rawat Publications*.

[101] Schoumaker, B. (2014). Quality and Consistency of DHS Fertility Estimates, 1990 to 2012 *DHS Methodological Reports; 12, ICF International: USA, 2014.* 106 p.

[102] Schoumaker, B., Dabire, H.B., Gnoumou-Thiombiano, B. and Key, J. (2006) Collecting Community Histories to Study the Determinants of Demographic Behaviour. *Population*, 61(1), 71-97.

[103] Shapiro, D. (2015). Accelerating Fertility Decline in Sub-Saharan Africa. *Population Horizons* 12.1 (2015): 3-12.

[104] Shapiro, D., Tambashe, B.O.(1997). Education, employment and fertility in Kinshasa and prospects for changes in reproductive behavior. *Population Research and Policy Review.* 16:259-287.

[105] Tertilt, M., Field, E., Molitor, V., and Schoonbroodt, A. (2016). Gender gaps in completed fertility. *Journal of Demographic Economics*, 82(2), 167-206.

[106] Tertilt, M., Doepke, M. (2018). Women's Empowerment, the Gender Gap in Desired Fertility, and Fertility Outcomes in Developing Countries. (No. 2018-001.)

[107] Towriss, C.A., and Timæus, I.M. (2017) Modeling period fertility: Schooling and intervals following a birth in Eastern Africa. *Population Studies*, DOI: 10.1080/00324728.2017.1370121

[108] United Nations (1995). Women's Education and Fertility Behavior: Recent Evidence from the Demographic and Health Surveys. *New York: United Nations*.

[109] Usher, D. (1997). Education as Deterrent to Crime. *Canadian Journal of Economics* 30 (2), 367- 384.

[110] Van Bavel, J. (2004a). Detecting Stopping and Spacing Behaviour in Historical Demography. A Critical Review of Methods. *Population (English edition)*, 59(1):117-128

Van Bavel, J. (2004b). Deliberate birth spacing before the fertility transition in Europe: Evidence from 19th century Belgium. *Population Studies*. 58, pp. 95-107.

[111] Van Bavel, J. and Kok, J. (2004). Birth Spacing in the Netherlands. The Effects of Family Composition, Occupation and Religion on Birth Intervals, 1820-1885. *European Journal of Population*. 20:119-140.

[112] Vogl, T.S. (2015) Differential fertility, human capital, and development. *The Review of Economic Studies*, 83(1), 365-401.

[113] Watkins, S.C. (1987). The fertility transition: Europe and Third World compared. *Sociological Forum*. 2(4):645-673.

[114] Western, B., Kling, J.R., and Weiman, D.F. (2001). The labor market consequences of incarceration. *Crime and delinquency*, 47(3), 410-427.

[115] Winikoff, B. (1983). The Effects of Birth Spacing on Child and Maternal Health. *Studies in Family Planning*. 14(10):231-245.

[116] Witte, A.D., Tauchen, H. (1994). Work and crime: an exploration using panel data. *Public Finance* 49, 155-167.

[117] Woitkievicz, R.A., McDonald, E. (1987). A Mimic Model for the Relationship Between Education and Fertility. *87 CDE working paper*.

[118] Yashiv, E. (2007). Labor search and matching in macroeconomics. *European Economic Review* 51.8: 1859-1895.