A Life-cycle model of human capital formation and educational choices in developing economies

Rosalia Vazquez-Alvarez [a] and Taghreed Adam [b]

This version: 12th March, 2008 (Preliminary)*

[a] *Swiss Institute for International Economics and Applied Economic Research (SIAW), University of St.Gallen, St.Gallen, Switzerland*

[b] *Health System Financing Department, World Health Organization (WHO), Geneva, Switzerland*

Swiss Institute for International Economics and Applied Economic Research

World Health Organization

University of St.Gallen

Abstract

In impoverished societies a shock to the household's resources often results in a decision to reduce the contemporaneous investment on children's education. Although such shocks may be transitory in nature they could lead to permanent long run effects in terms of reduced human capital for future generations thus further enhancing a state of deprivation. This paper models the long-run effects of productivity shocks using a micro-economic founded life-cycle model where investment uncertainty faced by heterogeneous households plays a crucial role with regards to optimal educational choices for their primary school age children. The dynamics in the structural model provide the basis for identification and estimation of endogenous human capital formation allowing for household's heterogeneity, a concept often missing from empirically implemented macro-models of growth. The empirical section illustrates the modelling strategy and estimation procedure using data from Tanzania (The Kagera Health and Development Survey, 1991-1994). The results suggest that primary school age orphans that attend school regularly can experience gains in human capital (i.e., productive capacity within the household) that are 6.2% higher than similar orphans in a non-schooling regime. Howerver, when one or both parents are present to bring up the children, the gains that result from attending school are small when we compared these to the human capital gains that result from intergenerational transfers,i.e., gains purely from knowledge transfers between generations.

Keywords: Life-cycle models, uncertainty under uninsured uncertainty, education, human capital formation, micro-economic estimation, policy evaluation.

JEL – Classification: D31, D91, J23, J24

---

*Taghreed Adam, Health System Financing Department, World Health Organization, Av. Appia 20, Geneva, CH-1201, Rosalia Vazquez-Alvarez, Swiss Institute for International Economics and Applied Economic Research (SIAW), University of St.Gallen, Bodanstrasse 8, St.Gallen, CH-9000. We are very grateful to David Evans, Javier Ferri, Markus Froelich, Michael Lechner and Obert Pimhidzai for helpful comments. We are also grateful for comments from participants in the 2007 Summer Economic Seminar Series of the African Economic Research Consortium (Nairobi, 2007) and to participants at the internal research seminar of the WHO (January, 2008, Geneva). Address for correspondence: SEW, University of St.Gallen, Varnbüelstrasse 14, St. Gallen, 9000, Switzerland. Emails: adamt@who.int, rosalia.vazquez-alvarez@unisg.ch

1 Introduction

The economic growth and the potential to grow of any nation depend on both the accumulated capital at levels and the dynamic processes that enhance such capital over the nation's lifetime. By capital we mean the endowments that underline the nation's capacity to produce, i.e., the interaction between human and physical capital. Since the seminal work by Grossman (1979,1999) it is well known that education and health play an interrelated role at determining the productive capacity of individuals, that is, individual's human capital formation that later defines the productive capacity in the economy. The basic Grossman model (Grossman, 1979) on human capital emphasises how individual's good health (or 'healthy time') determines the individual's capacity to produce conditional on some starting up level of education. In countries where the burden of disease due to risk factors such as unsafe sex, malnutrition, malaria or Tuberculosis is significantly high, the implications from Grossman's model is one of economic decay for two reasons, induced morbidity and accelerated mortality: low levels of 'healthy time' (morbidity) reduce the potential to produce consumption for current survival but at the same time, lower consumption today lowers the required investment needs in order to produce the necessary healthy time for the productivity needs in periods ahead.[1] The result is that too many individuals reach the minimum level of health required for optimal production (the point $H_{\min}$ under Grossman's interpretation) at an earlier stage of an individual's life cycle than would otherwise be the case. In turn, this leads to premature mortality that affects a significant portion of the productive population (i.e., prime age adults) with the possible loss in human capital resulting on economic collapse. The very instructive and insightful work by Bell, Devarajan and Gersbach (2003, 2006) provide a General Equilibrium model sets up that sets down the theoretical implications of adverse health conditions (in their case, the existence of an AIDS epidemic) on human capital formation and ultimately, its effect on economic growth. The model in Bell, Devarajan and Gersbach (2003, 2006) is innovative because it provides a direct link between adverse health conditions and human capital formation, while indirectly relating to the effect that health has on intergenerational human capital transfers. Their model is calibrated to the South-African economy with different scenarios regarding education and health policies. Their empirical findings suggest that ignoring the effect that adverse health (e.g., AIDS) have on intergenerational transfers of human capital (i.e., children's lost in learning due to adult's premature mortality) the effect of epidemics such as HIV/AIDS on economic growth are far underestimated. Whereas the modelling strategy by Bell, Devarajan and Gersbach (2003, 2006) captures important behavioural aspect that determine growth (i.e., intergenerational links), the model is one of aggregation and calibration with static implications where the dynamic effects from heterogeneous households is not accounted for.

This paper aims at providing a structural life-cycle model of educational choices under adverse health environments with endogenous human capital formation. Drawing from both Magnac and Robin (1991, 19996) and Lechner and Vazquez-Alvarez (2007) we define a behavioural model where optimizing households in developing economies select among alternative educational choices for their productive children. A productive child is one old enough to participate in the process of production and at the same time is old enough to benefit from start up education, a definition that may vary by country and region within a country. Households are heterogeneous with regards endowments that determine their start up level of human capital. Thus, heterogeneous households make consumption and investment (educational) decisions in an environment that implies uncertainty with regards to the future returns from current investment choices. Uncertainty is modelled as stochastic innovations that are idiosyncratic to the households and informs each unit about the state of the world. Although the shocks are stochastic and time-independent they affect the educational choices in the household and, therefore, have a permanent effect in the form on human capital accumulation. Positive shocks may lead to human capital formation that may act as an insurance against future risks thus enhancing education in periods ahead. The opposite may result when the stochastic shock leads to an adverse state of the world so that lower levels of human capital may perpetuate the effect of contemporanous adverse conditions in periods ahead.

The modelling strategy is implemented to the Kagera Health and Development Study Survey (1991-2004) that provides panel structure data for the North-West region in Tanzania. Using this data set we estimate key human capital parameters for distinct households according to number of parents present in the household and parental educational attainment. The estimation procedure is nonparametric in nature and relies in the existence of informative data. Our results suggest that orphan households – i.e., households where both parents have died prematurely and where an older sibling is the household head – benefit significantly from schooling. The reason for this is that compared to orphans who declare to attend school regularly experience human capital growth rates that are 6.2% higher than orphans in a non-schooling regime. On the other hand, in households where one or both parents are present the gains implied by intergenerational transfers – i.e., the net human capital gains obtained from knowledge transferred from parents to childrne – far outweights the human capital gains implied by formal schooling education. For example, when children live in in households where both parents are present they experience human capital gains over a two year period that equals 81.5% in the case that children are not sent to school, and 80.4% when children are sent to school. The difference of 1.1% – that would otherwise be adverse to the the schooling choice – is negligent and statistically insignificant. Similar differentials apply when comparing one parent households in alternative schooling regimes. Overall we find that the presence of parental input is what matters most. Thus, comparing human capital transfers between parents

[1] This interpretation on the effect of health on production is simply that of the dual effect that health plays in Grossman's set up: health is viewed both a good consumption and as an investment good.

and children between single parent households and two parent househols sugges that an extra parent can lead to as much as a 33% differential in the productive capacity of the child even if the child does not benefit from formal education. These evidences suggest that for developing economies, learning from parents remains the most important components in terms of human capital formation.

This paper is organized as follows. Section 2 presents an overview of the existing literature on models of growth and human capital under adverse health conditions. Section 3 develops a dynamic life-cycle model of educational choices with heterogeneous households and endogenous human capital formation. Section 4 describes the conditions for identification and the estimation procedure for the key human capital growth parameters: the estimation method draws from the reduced form specification driven by the structural set up. Section 5 describes the Kagera Health and Development Survey that will be the basis for the empirical estimates presented in Section 6. Section 7 concludes. The appendices provide further technical and quantitative material.

2 Human capital, health and education in the developing world

Economic agents make life-cycle choices knowling that the outcomes from such choices are often subject to uncertainty as result of stochastically changing environments. Uncertainty is in fact one of the most important reason that explains sub-optimal growth in developing economies, i.e., compared to the developed world, higher levels of policy uncertainty in the developing world often results in lower investment, lower output and consecuently lower rates of economic growth in the long run (see for example Barro, 1991, Aizenman and Marion, 1993, Alesina and Perotti, 1996 or Jeong, 2002, but to mention a few of the examples that explore the effect of uncertainty on investment choices in low income countries). But investment uncertainty – and its effects – is not a problem exclusive to policy makers. Households as independent units within the economy also face short run investment choices that are subject to uncertainty and, at the same time, these choices often determine the household's long run productive capacity. This is certainly the case in the developing world where insurance markets (e.g, insurance for health but also labour market insurance) remain underdeveloped.

One very important investment choices that households in developing economies have to face is that of educating the children in the household, i.e., in the presence of a weak social security system, adults in the households may think of investing on their children's education to enhance the future productive capacity of the household. But current investment on children's education implies forgone current consumption, both because the schooling choice may imply a monetary cost but also because the household forgoes the opportunity cost of education, i.e., the gains that primary school children can bring if these are sent to work instead. The effect of sub-optimal education in adverse health environments has already been captured by Bell, Devarajan and Gersbach (2003, 2006): their work models how parental subjective uncertainty on their children's survival affects the educational choice that these parents make for their children. More specifically, in the presence of an epidemic – e.g., high HIV-prevalence – the subjective expectations that parents hold regarding the survival of their children is significantly lower than in the absence of and epidemic. Such expectations become crutial at determining parental choices on their children education. A reason for this is that in the developing world the strength of the social family network – and the absence of a stable social secutiry system – implies that households invest on their children's education not just for altruistic reasons but also to secure the consumption needs of the household in the future, i.e., to secure consumption when current adults in the households reach the end of their productive lives. In the presence of adverse health conditions – e.g. when high HIV-prevalence and incidence of premature mortality from AIDS is common knowledge – parents will hold lower subjective survival rates on their children than would otherwise be the case. This now higher level of uncertainty lowers the future returns that parents would expect to obtain from contemporaneous investment on children's education, and more so where the opportunity cost of current investment on the education of children can be measured against the consumption gains obtained by the households if the primary school children – i.e., old enough to be productive – are sent to work instead.

Bell, Devarajan and Gersbach (2003, 2006) – BDG from now on – analyse the effects of sub-optimal educational investments on human capital formation; we understand by human capital the productive capacity of economic agents that ultimately determines long run economic growth. The BDG macro-economic model accommodates the two most important components that determine human capital accumulation, namely, human capital resulting from intergenerational transfers – i.e., the transfers of productive knowledge from parents to children – and human capital aquired as result of formal education – e.g., the effect of attending school regularly on the future productive capacity of the child. Using South African data at the macroeconomic level BDG calibrate their model to simulate the effect of no policy intervention – e.g. no education or no distribution of orphan children among households. South Africa is a particularly interesting example because – at present – it is estimated that the country suffers a 25% HIV-prevalence. The empirical findings from such simulation imply that with subjective probabilities that reflect the HIV/AIDS epidemic, the absence of interventions would lead to the collapse of the South African economy in less than four generations, i.e., in less than 40 years all households would experience negative income growth alongside zero schooling attendance for primary school age children: much of the negative outcome can be imputed to the broken link between generations (i.e, low intergenerational knowledge transfers) as result of parental premature mortality.

Wherease the BGD is fundamentally a macro-economic approach, its foundations are based on behaviour at the micro-economic level. That is, individual's – or household's – accumulation of human capital formation is the result of dynamically changing investment decisions on human capital formation that maximising agents make when choosing an optimal path over the life-cycle. At the the same time optimal choices depend on individual's preferences and characteristics that ultimately determine the individual's level of risk aversion, i.e., individual's reaction towards uncertainty when making optimal choices, for example, investment choices that shape their ability to produce in periods ahead. Households in developing economies can be thought as single units of production that make investment choices and these choices include their children's schooling regime; the choice is both for the benefit of the children but also for the overall benefit of the household. But likewise individuals, households are economic units that at the micro-economic level will behave idiosyncratically. This means that our understanding of long run economic growth can benefit from an understanding of the micro-economic foundations of human capital formation, especially if a micro-economic approach is able to integrate the main human capital formation concepts from BDG, that is, allowing for both the effect of investing on children's human capital formation and the effects of living in adverse health conditions that can undermine the health stocks of all household members. Grossman (1972, 1999) provides the framework to interpret the micro-economic effects of investment decisions on human capital while accomodating adverse health conditions – i.e., negative effects to present or future expected health stocks. According to the seminal work by Grossman (1972), health stocks enter the utility of individuals over the life-cycle both as goods that we produce when we invest in our health as well as goods that we consume in form of healthy (i.e., productive) time. Thus, according to Grossman's interpretation, productive human capital is the interaction between health stocks and knowledge stocks. In the developing world, the concept of the extended household implies that the productive capacity of 'surviving' children is in fact part of the adult's human capital stock when these adults are no longer productive, i.e, at some point it is expect that adults in the household are old enough so that their stocks of health produce zero healthy (productive) time for any given level of education (knowledge stock), and it is their children's stocks of both knowledge and health that act as human capital for these elderly household's members. This reason along can justify the need to invest on children's human capital – with regards to health but also with regards to intergenerational transfers and formal schooling – even in the absence of altruism. When parents hold low subjective survival probabilities on their children it may be optimal to shift any of the investment on their children's human capital formation (i.e., knowledge and health stocks) to alternative choices that imply lower risk, for example, increase the contemporaneous consumption capacity in the household by sending the children to work instead. Furthermore, Grossman's (1972) predictions can also explain why adverse health conditions that have a direct effect on the health stock of adults may translate into significantly lower human capital investment for the children in the household. When adults suffer the effect of some illness, morbidity increases the amount of time spent ill and lower the time spent in productive activities. This reduces the resources available in the household including fewer resources employed to enhance health stocks and less quality time spent with the children. But part of the investment on individual's health stock is based on investing on their children's health stock as means to survive into old age. Thus, parental morbidity can reduce the current health stock of children even if these do not suffer the illness directly. Moreover, lower time spent with children as consequence of morbidity can lead to less quantity of intergenerational knowledge transfers from parents to children. At the micro-economic level this results in children that grow up with sub-optimal human capital formation within the household. When the effects of morbidity are significant for the economy (e.g., in the presence of high HIV-prevalence) the aggregation of these micro-economic effects become significant for economic growth, as predicted in the work by BGD that makes use of macroeconomic foundations.

The macroeconomic based general equilibrium framework in BGD and its calibration approach provide a clear quantitative interpretation: the calibration process takes the dynamic changes on real gross domestic product as the key signal to measure the evolution of human capital formation in adverse health environments, for example, in South Africa where nowadays the HIV epidemic directly or indirectly affects at least ¼ of the prime age population. The problem with macroeconomic models is that these cannot take into account individual's idiosyncratic behaviour and how such behaviour affects human capital formation. Analysis at the aggregate level implies that all individuals in the population are homogenous with respect to the calibrating parameters, for example, with respect to risk attitudes and subjective survival probabilies. But nowadays the existence of fast computers alongside large data sets implies that it is possible to account for individual's heterogeneity to understand the investment decisions at the micro-economic level, for example, investment decisions that may result in human capital formation. If the available data is population representative the estimates can be used to infer from the sample to the population, while panel data sets provide the correct structure to analyse dynamic changes in the population without having to rely on aggregates and calibration. The advantage of using data at the microeconomic level is that it allows for the behaviour of individuals (i.e., observed preferences and characteristics) to underline the final estimate of human capital formation for heterogenous agents in the economy. For example, we can examine the characteristics and preferences of two parents versus single parent households to understand what determines the educational choices that parents make for their children and at household level. In general, the micro-economic approach requires the development of structural life-cycle models that capture the effects from individual's preferences and characteristics to the dynamic investment choices that lead towards human capital accumulation. Structural life-cycle models have already been developed with referece to optimal life-cycle labour market choices (e.g., Magnac and Robin, 1991, 1996, Browning, Hansen and Heckman, 1999, Costa-Dias, 2002, or Lechner and Vazquez-Alvarez, 2007). The framework by Magnac and Robin (1991, 1996) defines a dynamic behavioural problem where optimizing individuals choose among discrete labour market regimes namely, employment or self-employment. The theoretical and empirical work by Costa-Dias (2002) follows a similar framework where individuals choose between the status of employment or unemployment allowing for the stochastic labour market changes and idiosyncratic uninsurable uncertainty that determines the individual's fixed cost implied by the labour market regimes. Lechner and Vazquez-Alvarez (2007) further allow for a third labour market regime – that implied by the existence of Active Labour Market Policies (ALMP) – and the implication these have with respect to helping in the process of human capital formation.

The common concept in the works by Magnac and Robin (1991, 1996), Costa-Dias (2002) or Lechner and Vazquez-Alvarez (2007) is the development of a structural life-cycle models where net gains from productive activities (i.e., earnings) play a central part in signalling investment choices leading towards human capital formation. For example, in the paper by Lechner and Vazquez-Alvarez (2007) unemployment for a low skill individual is an optimal choice when comparing his benefits from unemployment to the earnings offered in the market for low skill labour. A sudden productivity shock – e.g., a wage subsidy, a sudden shortage of skill workers, etc. – can induce a change of regimes from unemployment to employment if such shock implies that the benefits from work are now higher than the receipts from unemployment benefits. In the new employment regime the individual retunrs have two components: the monetary gains from work that are now higher than the unemployment benefits but also the gains from work experience, that is, enhanced human capital in the learning by doing process that ultimately lead to higher earnings in periods ahead. Thus, the regime change from unemployement to employment is an investment decision that implies a risk – measures by the forgone unemployment benefits – but in the process of working the newly employed worker enhances his skills by accumulating 'knowledge' and it is this accumulation of human capital formation that implies a lower risk attached to to the working regime, that is, relative to unemployment it become less risky to choose employment in the future because enhanced human capital induces higher earings thus reducing the opportunity cost from unemployment in periods ahead (i.e., the monetary value of unemployment benefits). Based on such principle, the life-cycle models in Magnac and Robin (1991, 1996), Costa-Dias (2002) or Lechner and Vazquez-Alvarez (2007) develop structural life-cycle models leading to reduced form specifications that imply identification of key human capital parameters. Since human capital accumulation is the result of labour market participation, earnings are the key signal to understand investment decisions that lead to enhance productive capacity. Thus, the use of earnings in structural life-cycle models at the micro-economic level to understand human capital formation (e.g., as found in Costa-Dias, 2002, or Lechner and Vazquez-Alvarez, 2007) is analogous to the use of GDP to understand human capital formation within a dynamic macro-economic setting, for example, as found in the work by Bell, Devarajan and Gersbach (2003, 2006).

The basic structures used to understand the investment behaviour of productive agents in the developed world can also be used to construct models that help us capture the investment behaviour of agents in developing economies. Perhaps the crutial difference between the developed and the developing world is the definition of a productive unit. In the developed world the productive unit is the individual since children would never be considered as units capable of producing income for the purpose of consumption in the household. This is not the case in the developing world where children (e.g., old enough to attend school) can be though productive household units. Thus, in the developing world it is the households as a unit (and not the individual as a unit) that determines the productive capacity of individuals and children who are old enough to contribute to the process of production are also intrinsic parts of the household productive capacity. Other than this, agents (i.e., households in the developing world and individuals in the developed world) take net gains from productive activities as the central indicator for the agent's investment choices. In the case of the developing world, this involves investment choices that affect the school attendance of productive children. When adverse conditions – e.g., adverse health socks that affects the adults in the household – lead to decisions that imply a non-school regime for the children, the short run decision may imply higher contemporanous consumption but at the expense of forgone investment on the long run accumulation of human capital for the households. This in turn leaves the household in a weaker position to deal with future adverse productivity and health shocks, thus implying the potential for policy intervention, for example, intervention that subsidized schooling but also other interventions that prevent the effect of productivity and health shocks.

One way to understand the effects from policy interventions is to quantify the effect that the household's investment decision making have on the household's human capital accumulation. The work by Costa-Dias (2002) and Lechner and Vazquez-Alvarez (2007) draw from Heckman, Lohner and Taber (1998, 1999) to quantify endogenous human capital formation that is different from experience. Heckman, Lohner and Taber (1998, 1999) argue that observed distributional difference in the returns from labour market participation result because human capital – i.e., the potential to produce in the labour market – differs in unobservable ways between participants that otherwise have identical observed labour market characteristics. Thus, the skill price may be identical for all participants sharing the same skill class, but the individual specific idiosyncratic capacity to produce in the work place implies 'augmenting' such basic skill price (i.e., the rental price of labour) in specific ways for each of the labour market participants. The econometrician observes the final returns and measures the 'common to all' basic rental skill price augmented by the individual specific human capital. According to Heckman, Lohner and Taber (1998, 1999), individuals can invest to enhance their own human capital (e.g., on the job training) at the expense of forgone income since there is rivalry between working and learnings. This argument implies that human capital depends on education, individual's idiosyncratic characteristics (e.g., ability) and labour market participation, i.e., human capital formation is endogenous and differt from experience and it ought to be measured as an augmenting parameter that determines the idiosyncratic capacity that individuals have to earn a return from active participation in the process of production.

The key concept in the work by Heckman, Lochner and Taber (1998, 1999) is to think that investment in learning shapes the formation in human capital conditional on individual's characteristics. We draw from such principle as well as follow the work by Costa-Dias (2002) and Lechner and Vazquez-

Alvarez (2007) to develop a life-cycle model of discrete educational choices with endogenous human capital formation at the household level. We assume that heterogenous households decide about sending their children to school, a decision that implies investment under uncertainty and it is this uncertainty that prevents or induces the actual schooling choice. In fact only uncertainty can explain the dynamic educational choices in the household: in its absence the rational risk avers household makes an optimal schooling choice (attendance or not) for the full lifetime of the household. In the presence of uncertainty risk avers households update their educational choices after the uncertainty is realized at each point in time and conditional on expectation about realization of stochastic shocks in the future. However, uncertainty does not reflect price and income volatility (e.g., as is the case in Meguir and Pistaferri, 2001). Instead, we follow from Lechner and Vazquez-Alvarez (2007) to make similar assumptions as in Magnac and Robin (1991) that likewise follow the work by Kihlstrom and Laffont (1979). In doing so we assume that household's uncertainty about future returns on their children's schooling is explain fully in terms of attitudes towards risk, where the concept of risk is measured as 'the opportunity cost' implied by alternative investment regimes. To model this we assume that each household has some latent or hidden reservation value for each possible schooling choice (e.g., send the child to school or not). The reservation value depends on the fixed cost that each regime may imply. The data does not directly reveal the fixed costs for each of the schooling regimes but we know that such fixed costs are idiosyncratic to the household and depend on the household's preferences and characteristics, variables that are in fact observed in the data. When households are observed to send their children to school they signal that relative to the state of the world the schooling choice implies reservation value sufficiently low relative to the productivity effects implied by an stochastic changing world, that is, conditional on preferences and characteristics the household invests on human capital formation because the state of the world is favourable. The result in that human capital in the households is enhanced beyond that implied by intergenerational transfers. Alternatively, when a household receives a negative shock (i.e., the state of the world reveals negative effects on the household productivity process) the choice may be that of not sending the children to school, and the result may be a loss in human capital formation for the the children but also for the household.

Our proposed life-cycle model aims at providing a structural model that captures all arguments just mention. This implies evaluating the impact of schooling investment decision allowing for 'earnings' to signal human capital formation. The unit of interest that defines earnings and investment decision is the household although it is 'the children' in the household that provide the identification means for human capital formation. We allow for households in developing economies to be heterogeneous as result of adverse health conditions that determine different household types, an assumption that follows closely the set up in BDG but within a micro-economic framework. The final aim is to use the structural set up to define a set of necessary condtions to identify an estimate the key human capital parameters, i.e., growth rates in productive activity that may result from the combination of learning as result of intergenerational human capital transfers and learning as result of investment on formal schooling. The empirical implementation of the model aims at providing policy evidence on the effectiveness of particular interventions (e.g., education, health knowledge) on the process of human capital formation that ultimately determines the growth rate in the economy.

3 A life-cycle model of educational choices and human capital formation

This paper aims at developing and implementing a life-cycle model of educational choices and endogenous human capital formation allowing for heterogeneous households with respect to start up endowments. To this aim we take *the household* – and not the individual – as the unit of interest. We assume that the representative household i is born at time $t=0$, the moment when such unit is formed and starts to participate in the productive activities of the economy. The representative household lives for a fixed number of discrete time periods T. At each point in time $t \in [0,T]$ the household evaluates the state of nature and decides optimally about spending on current consumption and education for the children in the household: although contemporaneous educational choices have only a *current* direct effect on children's human capital, the choice to spend on education should be viewed as an investment choice that defines the future human capital formation for the household in periods ahead. At the same time, children in the household are viewed as productive units that may be less efficient than parents at turning labour-time into consumption goods but are potentially able to (inelastically) supply as much time towards productive (non-educational) activities as adults.[2] However, unlike adults, children's activities can also take the form of schooling, an activity that may be thought as rivalling the time that children may devoted to the production of current consumption. In our model we use a simplifying assumption so that the choice between schooling and working for children is mutually exclusive and identical for all children in the household who are of sufficient age to enjoy the effect of schooling. Thus we define the binary indicator e_{it} such that $e_{it}=1$ if the ith representative household chooses to send the household's children to school at time t, and $e_{it}=0$ otherwise.[3]

[2] In developing economies it is more realistic to assume an inelastic supply of labour with the cost of leisure close to zero, an assumption consistent with labour markets where there is no insurance against unemployment.

[3] A more generalized form would imply letting $e_t \in [0,1]$ be the fraction of time devoted to education so that on average each child in the households supply $(1-e_{it})$ time units of labour as in the theoretical setting of Bell, Devarajan and Gersbach (2003, 2006). We choose a discrete interpretation consistent with the empirical evidence that often can only detect a binary process with respect to school attendance. Our assumption is consistent with the possibility that children participate in household production activities as long as these do not preclude school attendance. Theoretically we are giving all children the same household treatment irrespective of gender and origin (e.g., biologic offspring or adopted child) as long as the child is sufficiently old to enter primary school.

The endowments with which households are born at time $t=0$ determine the ith household's type s_i and with this the first level of heterogeneity in our model. The ith representative households is born with $n_{it=0}>0$ number of *productive* children while the number of adults in the household ($m_i \geq 0$) differs according to the household's history of adverse health events with different number of adults commanding different start up levels of human capital, i.e., productive capacity that turns into income in the form of consumption goods. Let $s_{it=0}=s_i \in \{s_1,s_2,s_3,s_4\}$ denote the realized household type at time $t=0$ such that s_1 defines a household where both parents are present, s_2 defines households where the mother dies prematurely, s_3 defines households where the father dies prematurely and finally s_4 defines households with orphans (where both parents die prematurely). Thus, at $t=0$ the ith representative household has a total of ($m_i+n_{it=0}$) individuals all of which have the potential to contribute towards the production of consumption goods.[4] Whereas household's formation is exogenous to the model, the number of household children n_{it} changes endogenously over the household's life-cycle as result of both fertility and the endogenous accumulation of human capital. This also means that there is a difference between the *countable* number of children (n_{it}) and actual value (benefit) that the household gets from having n_{it} productive kids. For this reason we introduce the variable $\tilde{n}_{it}$ that defines the number of productive children in the household measured in terms of their potential in the process of production and human capital accumulation.[5] For any given household with $m_i+\tilde{n}_{it=0}$ productive members, starting up household's endowments in the form of start up human capital is characterized by the number of adults in the household at $t=0$. The number of adults is assumed to remain fixed over the household's finite lifetime such that $m_{it=0}=m_i$, $\forall t \in [0,T]$.

This theoretical simplification does not affect the qualitative nature of the structural model and can be relaxed at the empirical level to quantify heterogeneous schooling treatments among children in the same household.

[4] Younger children are non-productive children and consume a negligent amount before reaching the point at which educational choice have to be made for them. Likewise, older household members that cannot contribute towards the productive process are assumed to consume a negligent amount.

[5] That is, $\tilde{n}_{it}$ quantifies the number of children augmented by the human capital benefit to the ith household of having n_{it} children at t. To clarify this, let there be two households (i,l) such that $n_{it}=n_{it+1}$ and $n_{lt}=n_{lt+1}$. In both cases the growth rate as result of fertility is zero. However if household i 's human capital is of better quality than that in household l, and if such human capital is transferred to the children as a household's public good, then the n_{it} children at time $t+1$ have grown in terms of productive value more than the n_{lt} children at $t+1$: in each case we allow for $\tilde{n}_{it}$ and $\tilde{n}_{lt}$ to interpret both physical growth as result of fertility and the additional change (positive or negative) as result of changing (increasing or decreasing) human capital in the household. We may characterize the example by assuming that human capital includes better health knowledge. Thus, comparing the growth in terms of productive value between $\Delta \ln \tilde{n}_{it}$ and $\Delta \ln \tilde{n}_{lt}$ quantifies the value of a higher probability of children's survival into adulthood of n_t for any of the two households. This assumption is not too dissimilar to that applied in labour economics between actual observed 'years of experience' for a given skill class and actual 'human capital' that varies among individuals with otherwise identical experience and skill class.

Each child in the economy is associated with a unique $s_i \in \{s_1, s_2, s_3, s_4\}$ type that determines the start up level of human capital $\Lambda_t(s_i)$ as result of the number of parents present in the household. Whereas each child in a given household (and therefore the household) inherits $\Lambda_t(s_i)$ exogenously, sending children to school may lead to an endogenous increase in the household's human capital as result of learning. That is, $\Lambda_t(s_i)$ can change endogenously over the finite horizon $t \in [0, T]$. However, the choice to educate the children in the household implies a current investment on human capital that becomes only productive in future periods at the expense of reducing contemporaneous household production by an amount equal to that of children's productive capacity. We can assume that all production is used for consumption where consumption c_t is a homogenous good shared equally among all household members. Because education is only productive in periods ahead (or the long run) we assume that households derive current utility only from current consumption, that is, a rational forward looking household maximises an inter-temporal utility function with a single argument c_t summarizing the source of utility for a given educational choices.

We have defined the first level of heterogeneity according to exogenously determined household type s_i where $s_i \in \{s_1, s_2, s_3, s_4\}$ underlines starting up human capital endowment $\Lambda_0(s_i)$. We can interpret these as the ability of the household to transfer productive skills to their children where the child-rearing capacity from adults is assumed to be monotonically increasing in the number of adults present in the household, that is, $\Lambda_0(s_i) \geq \Lambda_0(s_l)$ for $s_i \leq s_l$, $(i, l) \in \{1, 2, 3, 4\}$. However, not all households with the same number of adults are equally effective at transferring knowledge to their children since transfers can differ according to observed variability (e.g. education) as well as due to unobserved variability (e.g., innate ability, motivation, etc). Let θ_i define the household's skill specific ability to transfer human capital onto future generations. In terms of observed variability we can interpret θ_i as the average educational level of adults in the household.[6] With this we modify the representation of the household's start up human capital as $\Lambda_0\left(s_i(\theta)\right) \geq \Lambda_0\left(s_l(\theta)\right)$ for $s_i\left(\theta\right) \leq s_l\left(\theta\right)$, $(i, l) \in \{1, 2, 3, 4\}$, where θ is the irreversible average schooling achievement of adults at $t = 0$. Simplifying $s_i \in \{s_1, s_2, s_3, s_4\}$ to be s_j for any *ith* household where $j \in \{1, 2, 3, 4\}$, it is assumed that among households with identical s_j characteristics, better skills leads to higher start up levels of human capital such that $\Lambda_0\left(s_j(\theta^*)\right) \geq \Lambda_0\left(s_j(\theta^{**})\right)$ for $\theta^* \geq \theta^{**}$ where $(\theta^*, \theta^{**}) \in \Theta$ and Θ is the space containing all possible

[6] The parameter θ_i is likely to vary significantly between households with identical educational level as result of innate ability, motivation and other variables that are rarely identified by the data. To let θ_i interpret the average

ability values in the population. For simplicity we assume that θ arrives exogenously and may even have the capacity to determine s_j type.[7] Furthermore, it is assumed that θ remains unchanged throughout the life of the household. Together the assumption of fixed $m_{it=0} (= m_{it} = m_i)$ and fixed θ implies that the household's human capital can only change as result of a combination of educating the household's children and any endogenous change that increases both the number and quality of the children in the household: the assumption rules out human capital growth as result of enhancing the adult's start up level of education.

Our third and final level of heterogeneity refers to the existence of uninsurable uncertainty. In the presence of alternative productive activities, a choice to educate the children in the household implies taking a decision under uncertainty that either prevents or induces the actual choice. The main reason for such an uncertainty is the presence of epidemics, wars and other socio-economic and productivity factors that become of common knowledge to the population (including the household) and results in the probability of children surviving into adulthood to be significantly less than one. Under such survival probability the opportunity cost of investing in education is priced and high. This is true even if education is free from pecuniary costs in environments where children's work is accepted so that the opportunity cost of going to school is clearly defined according to the children's productive capacity. We model household's uncertainty by assuming that the state of nature changes stochastically at each and every time over the household's finite life-time. The stochastic change is introduced as a time-independent idiosyncratic shock π_{it} that will determine how much *effort* the household is willing to employ in educating the household's children. The stochastic shock π_{it} is transitory in nature (i.e., the state of nature changes stochastically at each point in time) but nevertheless the household's reaction to the innovation leads to a permanent effect through future changes in the form of human capital (accumulation or depreciation). To illustrate the effect of a shock π_{it} we start from a situation where all members in household i – and this includes the productive children in the household – contribute towards producing consumption goods at t so that $e_{it} = 0$: this choice may be optimal if the household's subjective probability on their children's survival into adulthood is sufficiently low such that $e_{it} = 0$ implies the low risk alternative relative to the risky choice $e_{it} = 1$. Let π_{it+1} be an innovation that acts as a positive shock to the productive capacity in the household, for example, a sudden farm subsidy that increases production or in fact leaves household production unchanged in the event that children are sent to school at $t+1$, that

educational achievement implies letting θ_i take discrete values (e.g., below primary education, up to secondary education and beyond secondary education).

[7] For example, Fuchs (2001), Gruber (2005), Froelich and Vazquez-Alvarez (2007) but to mention a few, would provide theoretical constructs and empirical evidence to suggest a direct link between better education and lower chances of premature death, for example, by means of a more effective use of information on nutrition, reduce risk behaviour with respect to HIV/AIDS and generally better use of preventive methods to reduce mortality.

is, $e_{it+1} = 1$ becomes optimal once the stochastic shock π_{it+1} is realized. The choice $e_{it+1} = 1$ results as consequence of a *transitory* shock π_{it+1} but it leads to a *permanent* effect in terms of increased human capital for the household in periods ahead through two channels; first through the direct impact that increased children's knowledge has on the knowledge in the household, and secondly because more educated children have a better chance to survive into adults thus further enhancing the present level of productive capacity in the household via the enhanced value that levels of human capital have on existing and future productive children.[8] As the choice the educate the children in the household is made, the alternative choice of no education in periods ahead ($e_{i\tau} = 0,\ \tau > t+1$) becomes less likely, thus reflecting the fact that as more effort is put into education the more that education becomes the optimal choice in future periods. To a large extend allowing for an stochastic shock π_{it} reflect the fact that educational choices made in adverse environments are potentially risky decisions with productivity choices (educational choices included) made by agents (in this case households) whose preferences display risk aversion. Clearly, a schooling choice reduces the probability that children die prematurely[9] and further helps to accumulate household knowledge (human capital) thus further insuring the household against future (potentially adverse) shocks.

In our modelling strategy the introduction of heterogeneous households allows for the identification and estimation of human capital growth resulting from dynamic optimal choices under stochastic life-cycle shocks. These shocks ultimately affect the micro-economic educational attainment of individuals with final human capital growth effects at the macro-economic level. The approach to endogenous growth formation with heterogeneous agents follows in the line of work such as that of Bell, Blundell and Van Reenen (1999), Browning, Hansen and Heckman (1999), Blundell, Costa-Dias and Meguir (2002) and Lechner and Vazquez-Alvarez (2007), all of these being examples closely related to evaluation studies of labour market programmes and policies in the developed world. In tune with all the aforementioned studies we also develop a life-cycle model with micro-economic foundations that contrasts with macro-economic models of endogenous growth where aggregation inevitably leads to the elimination of the effect that individual's preferences at the micro-economic level may have on optimal choices and human capital formation. We present the model by first describing the household's objective function and optimal decision rules that become crucial at defining the dynamics of the state variables. In developing the model we show that conditional on individual household's characteristics – determinants of

[8] We assume that knowledge or other building blocks of household's human capital are share equally among all children in the household at time t and inheriting it as a public good. Since human capital is cumulative it will also enhance the worthiness of future children (unburned) or children that are yet too young to be defined as productive.

[9] This is not just as result of children's increased knowledge on issues such as, for example, HIV/AIDS transmission but also because attending school effectively reduces other risk factors associated with children's premature

their preferences and therefore level of risk aversion – and together with the idiosyncratic innovations disclosed in each period, endogenously defined human capital determines optimal educational choices.

3.1 The household's objective and educational choices

It is assumed that ith representative household starts its life at time $t = 0$ with $m_i \geq 0$ number of productive adults and $n_{it=0} \geq 1$ number of productive children subjectively valued as $\tilde{n}_{it=0} \geq 1$ as result of the household's start up human capital at levels. The formation of the household is exogenous to the model and determines start up human capital conditional on skill class θ_i. The household lives a finite life defined by the interval $[0, T]$ implying that adults in the household retire at $T < \infty$ with some fixed level of human capital. Allowing for children to be part of the household does not imply an extended dynasty. Instead, we think of the household as a purely productive unit with concave human capital that increases from its initial start up level and as result of the children's human capital formation where the latter is though as a public good for the household. At time T the children in the household reach the flat portion of the human capital curve at which point separation occurs in order to form a new household. [10]. Optimal allocation of lifetime resources implies maximizing an objective function at each $t \in [0, T]$ over the entire household's working horizon. The household's objective is summarized with a separable time-varying utility function $u_{it} = u_t(c_{it})$ where consumption at t (i.e., c_{it}) is the only argument from which households derive utility. Indexing the consumption bundle with (i) implies difference in consumption between households as result of difference in number individuals in the household (even if households share the same type and skill class) while the time index (t) suggests that the consumption needs of the household can differ over time.[11] We model utility so that it can be time-varying (i.e., $u_t(.)$) but it is assumed to be homogenous between households, that is, utility is heterogeneous between households only as result of different consumption needs. Finally we notice that leisure is excluded from utility despite the fact that

mortality (see for example Lopez et al (2006) for a detail overview of the key Risk Factors associated with the Burden of Disease in developing economies).

[10] Our aim is to identify human capital growth resulting from educational choices conditional on individual household's characteristics. Monotonic growth implies that increments in human capital becomes zero if the change in the household's productive capacity is zero over any two consecutive time periods, an event that may occur when the household retires in the sense that the children go away to form new households and human capital in the old household remains constant (at levels). Thus, because we focus on rates of human capital growth we need not equate T with the death of the household but simply with the cessation of increasing human capital growth that occurs when the children in the household have accomplished their full human capital potential and they separate to exogenously form a new household.

[11] For example, in Bell, Devarajan and Gersbach (2003) household consumption is defined as $(m_t + n_t\beta)c_t$ for any given ith-household where c_t is a homogenous level of adult consumption and β adjusts c_t to represent children's consumption relative to that of adults. Our life-cycle model simply requires that we distinguish between heterogeneous households thus we generalize $(m_t + n_t\beta)c_t = c_{it}$ where the household suffix summarizes different consumption needs according to household size that includes consumption of non-productive members.

ours is a model that effectively emulates the utility that agents obtain from labour supply. Later (see Section 3.6 and Appendix 1) it becomes clear that excluding leisure from the utility function helps to identify a unique solution that characterizes the representative household's optimal choice. The exclusion, however, is well justified because as it eliminates wealth effects from the structural model, effects that – under homogenous preferences – would only be relevant for households at the top of the wealth distribution.[12]

We emphasis that our structural model aims at capturing the behaviour of individuals that choose optimally among mutually exclusive educational investment alternatives where only fixed cost from schooling and uncertainty about future returns from such choices causes low school attendance. Thus, incentives to make children attend schools should emphasis such fixed cost of attending school (e.g., reservation price for school attendance in the form of children's production) while assuming that there is sufficient infrastructure to absorb the required demand for education in the population. This is therefore a framework that will help us analyse socio-economic policies – e.g., schooling subsidies, employment creation, health subsidies, etc. – that directly target the behaviour of agents (the household) and not the supply side of the economy.[13]

So far we have all the ingredients required to determine the representative household's dynamic choices that change according to changes in the state of nature. At each time period t over the finite horizon the state space faced by the ith representative household is described by the amount of assets thus far accrued by the household (a_{it}), the household type (s_j) conditioned on skill & ability type (i.e., $s_j(\theta)$) that determines start up human capital unique to the household, the amount of human capital (Λ_{it}), the number and value of productive children in the household ($\tilde{n}_{it}$) and the idiosyncratic labour supply shock (π_{it}). Altogether these elements describe the state space faced by household i at time t that we define as $X_{it} = \left(s_j(\theta), a_{it}, \Lambda_{it}, \tilde{n}_{it}, \pi_{it}\right)$. We notice that except for $s_j(\theta)$ all elements vary over time. However, whereas a, Λ and $\tilde{n}$ are all assumed to change endogenously, the shock π is modelled as a

[12] Allowing for leisure to enter as argument in the utility function (i.e., $u_{it} = u_t(c_{it}, l_{it})$) opens the possibility that individuals who become wealthier as result of increased human capital can start to exchange work for leisure. This leads to the classic backward bending supply function associated with wealth effects taking over income effects. The complication in dealing with such functions is the possibility of multiple equilibriums for the solution of the value function (see Section 3 and Appendix 1). However, we are dealing with developing economies where labour supply is completely inelastic and therefore these are populations rarely associated by wealth effects. Thus, our assumption is well justified as would be the case for low skill populations in the developed word (see Lechner and Vazquez-Alvarez, 2007, for a similar assumption in the case of labour market policies for low skill individuals).

[13] This assumption is similar to that of assuming vacancy existence in the economy and thus voluntary unemployment in a framework where the key issue may be to analyze worker's behaviour and their reaction to active labour market policies.

random draw that we assume to come from some known distribution characterised by *iid* realizations.[14] Once the state of the world is well defined, the problem for the representative household of type $s_j(\theta)$ is one of maximizing expected discounted utility over a finite horizon as described in the following dynamic objective set up:

$$\begin{aligned}
&\max_{\{c,e\}_{t=0}^{T}} E_t \sum_{t=0}^{T} \left[\beta^t u_t(c_{it}) \middle| (X_{it}, \psi_t)_{t=0}^{T} \right]; \\
&where, \qquad X_{it} = \left(s_j(\theta), a_{it}, \Lambda_{it}, \tilde{n}_{it}, \pi_{it} \right) \\
&\qquad\qquad \psi_t = (r_t, W_{s_j(1)t}, ..., W_{s_j(\theta)t}, ..., W_{s_j(\Theta)t}, \Upsilon_{s_j(1,n)t}, ..., \Upsilon_{s_j(\theta,n)t}, ..., \Upsilon_{s_j(\Theta,n)t}, \tau_t) \\
&\qquad\qquad for \quad j = \{1,2,3,4\} \quad and \quad n \in \mathbb{N}^+
\end{aligned} \tag{1}$$

In expression (1) the constant β represents the discount factor and ψ_t is a vector describing the sequence of real prices that individuals face from period t to T : $W_{s_j(\theta)t}$ represents the monetary return from productive activities for adults in household of type s_j with skill $\theta \in \Theta$ so that $W_{s_j(\theta)t}$ is common to all i households with identical class (type) characteristics. The price $\Upsilon_{s_j(\theta,n)t}$ sums up the pecuniary costs of educating the children and can be equated to the productive capacity of children in the household who are sufficiently old to go to school: it is an amount assumed to depend on household's type ($s_j(\theta)$) so that $\Upsilon_{s_j(\theta,n)t}$ can be thought as the price offered by some employer to the parents of type $s_j(\theta)$ that choose to take their n productive children to work with them. At a macroeconomic level, $\Upsilon_{s_j(\theta,n)t}$ is common knowledge to all households in the population and may vary deterministically according to how many children ($n > 0$) are provided by the household. At a microeconomic level $\Upsilon_{s_j(\theta,n)t}$ varies by household according to the number of children in the household (n_{it}) and not the actual value that children imply to the household $(\tilde{n}_{it})$. The reason why $\Upsilon_{s_j(\theta,n)t}$ varies with n_{it} and not $\tilde{n}_{it}$ is because an employer pays a price for each productive child irrespective of the household's private valuation of the child. The number of population cells defined by $s_j(\theta)$ is finite and well defined at $t = 0$ for all households. The variable r_t represents returns from savings and τ_t summarizes net tax and transfers also at t : we can generalize $r_t \in [0,1)$ and $\tau_t \in [0,1)$ without any loss to the qualitative structural model. The vector ψ_t differs from

[14] The assumption that π is stochastic can be relaxed by assuming some dynamic process such that the shock becomes endogenous within the structural model like any other time dependent component of the state space. The assumption that these draws are independent and identically distributed implies that stochastic changes in the state of the world for a given household do not affect the state of the world of other households (independence) and all draws come from one single – assumed known – distribution (thus, identically distributed 'luck' for any given household in the population). However, how the household reacts to these random draws differs according to the household's characteristics (e.g., according to preferences, household's endowments or subjective believes).

X_{it} in that we assume perfect foresight for prices, that is, from time $t=0$ to T it is assumed that ψ_t is common knowledge to all households and this common knowledge is part of the information set efficiently used by households to make optimal life-cycle choices. Expression (1) explains the household's expected life-time utility conditional on the realization of the stochastic shock π that determines the state of nature: once the shock π_{it} is realized the uncertainty is revealed and the risk avers household makes optimal choices about consumption c_{it} and education e_{it}. The suffix t in $E_t(\cdot)$ can equally be substituted by π with expectations explained as $E_\pi(\cdot)$; this would emphasise that uncertainty refers only to the realization of the transitory productivity shock with ultimately permanent effects on growth rates of human capital. It is now necessary to describe the dynamics governing the state space that define the link between optimal educational choices and accumulation of human capital for the household.

3.2 The dynamics of assets, fertility and human capital

Household's lifetime optimal choices are subject to a set of constrains that evolve endogenously over time. According to the state space previously define we consider a problem constrained by three endogenously changing variables, namely, the number and value of productive children in the household ($\tilde{n}_{it}$), household assets (a_{it}) and human capital (Λ_{it}). We start by writing the law of motion that describes the accumulation of household assets where it is assume that these take into account both physical wealth and human capital that in turn depends on number of productive children:

$$a_{it+1}=(1+r_t)a_{it}+\left(e_{it}\right)\left\{\left(1-\tau_t\right)\tilde{W}_{s_j(\theta)t}\pi_{it}\Lambda_{it}\right\}+\left(1-e_{it}\right)\left\{\left(1-\tau_t\right)\left(\tilde{W}_{s_j(\theta)t}\pi_{it}+\tilde{\Upsilon}_{s_{ji}(\theta)t}\right)\Lambda_{it}\right\}-c_{it}\geq 0$$

$$\begin{aligned} &\textit{where} \quad && e_{it}\in\{0,1\}, && (2)\\ &\textit{and} && \left(1-\tau_t\right)\tilde{W}_{s_j(\theta)t}\Lambda_{it}\pi_{it}=W_{s_j(\theta)t}\pi_{it}\Lambda_{it} &&\\ & && \left(1-\tau_t\right)\left(\tilde{W}_{s_j(\theta)t}\pi_{it}+\tilde{\Upsilon}_{s_{ji}(\theta)t}\right)\Lambda_{it}=\left(W_{s_j(\theta)t}\pi_{it}+\Upsilon_{s_{ji}(\theta)t}\right)\Lambda_{it} && \end{aligned}$$

The positive inequality in expression (2) implies that the household's full wealth may be totally consumed at any current period but the same inequality allows for households to save and borrow so that we assume no liquidity constrains.[15] Accumulated wealth one period ahead is defined by savings (a_{it}) net from consumption (c_{it}) and augmented by income gains brought about by current production netted out from tax and transfers, i.e., $(1-\tau_t)$. Net income gains differ according to the household's contemporaneous

[15] The model allows for a lack of financial markets since $r_t=0$ is admissible. In the absence of financial markets households always meet consumption even if this implies running down assets or starving. In all cases it must be the case that $a_{t+1}\geq 0$ at any time period.

educational choices as described by the binary indicator $e_{it} \in \{0,1\}$. In households where children are observed to be educated at t (i.e., $e_{it} = 1$) net income gains depend on the returns $W_{s_j(\theta)}$ obtained from each unit of human capital commanded by the household at time t (i.e., Λ_{it}). In households where a decision is made to educate the children at t so that $e_{it} = 1$, it is the case that total returns from work – for a given level of human capital Λ_{it} – are augmented by the idiosyncratic transitory productivity shock π_{it} received at t. In households where productive children do not go to school (thus, $e_{it} = 0$) the net returns from productive activities equals the total returns ($W_{s_j(\theta)t}\pi_{it} + \Upsilon_{s_{ji}(\theta)t}$) obtained from each unit of human capital (Λ_{it}) where Υ_{it} explains the complement obtained from the working activities of n_{it} children; thus, $\Upsilon_{s_j(\theta,n)}$ is no specifically defined as $\Upsilon_{s_{ji}(\theta)}$ for the ith household with n_{it} children. The complement Υ_{it} implies an increase in the price received from each Λ_{it} unit of human capital commanded by the household but we assume that Υ_{it} is not affected by the productivity shock π_{it} since this 'productivity gain or loss' has only a direct effect on the decision makers in the household.[16] It is important to notice that the price $W_{s_j(\theta)}$ is not indexed by i because households of identical type are paid an equal amount for their skills ; in this respect we define $W_{s_j(\theta)}$ analogously as the so call 'skill rental price' put forward by Heckman, Taber and Lochner (1998, 1999). The dynamics of assets in (2) clearly show that human capital formation (i.e., a future increase in Λ_{it}) plays a key part in terms of future household assets (a_{t+1}) while these depend on the discrete educational choices ($e_{it} \in \{0,1\}$) made at each point in time.

Human capital accumulates as result of two interacting forces: first, the direct knowledge received from attending formal education that empowers children with human capital that further enhances the intergenerational knowledge transfer from their parents: it is assumed that schooling is a more powerful tool to enhance human capital formation than just the experience that working implies. Second, the existence of children in the household and their survival also has a positive effect on the production of household's human capital so that we can think of children per se as an investment source for the household human capital formation; the reason for this is that parental knowledge gets transferred to the children and, therefore, the existence of the children per se implies an enhancement of human capital in

---

[16] In households where children go to school ($e_{it} = 1$) only adults (decision makers) work and their returns are $W_t\pi_t\Lambda_t$. Adult's returns are equally affected by productivity shocks in households where children do not go to school; in these cases the household gets some benefit level from the working of the children that depends on the number of children but not on the productivity shocks that affects parents and parental schooling decisions. In this respect we are emulating a system of benefits within societies without unemployment but where children's contribution to the household production often act as some form of insurance when adverse conditions arise. As is the case with 'unemployment insurance' in the developed world, children's contributions to household's

the future. All households begin their life at $t = 0$ with a start up level of human capital determined by household's type $s_j(\theta)$ where $j \in \{1,2,3,4\}$ explains the number of adults in the household and $\theta \in \Theta = \{1,2,...,\theta,...,\overline{\theta}\}$ describes the adult's skill and ability level: empirically we cannot observe innate ability and motivation but assume that $\theta \in \Theta$ can be captured by the average educational attainment of the adults in the household. Whereas the human capital in the household may change endogenously as result of children's activities and existence, adult's start up level of human capital remains fixed throughout the household's lifetime and is fully defined by $s_j(\theta)$.[17] The accumulation of human capital as result of investing on children's formal education is assumed to be at a rate that depends on the type of household, the number and value of children in the household and the previously accumulated level of human capital. Furthermore, a positive productivity shock (e.g., a high value for π_{it}) leading to a choice such that $e_{it} = 1$ also acts as an augmenting parameter for the human capital in the household as result of the positive effect that better production (i.e., more resources) may have on accruing higher knowledge; Thus, the rate of human capital accumulation is augmented by the shock π_{it} so that a transitory change that is specific to each household leads to a permanent effect on household's human capital accumulation. However, even when the productivity shock is adverse, the existing human capital at levels also has a direct impact on the survival of the children in the household.[18] Although our modelling strategy does not explicitly assume human capital depreciation, an explicit interpretation of children's household growth as part of human capital can lead to depreciating human capital as result of childhood mortality but not as loss of human capital that once acquired will remain for as long as the household remains alive. The following expression defines the evolution of human capital with endogenously changing number of productive children that includes all the implications thus far discussed:

consumption in the developing world should invoke a *fraction* of earnings that do not depend on but do not imply productivity gains (as implied by π).

[17] The assumption is in tune with the empirical observation that there are distributional differences in earnings among household with identical s_j and θ characteristics. Endogenous human capital accumulation explains such differences in the same manner as differential in earnings among similar skill class is explained by individual's heterogeneity with regards to labour market experience and unobserved effort.

[18] We may justify the assumption in various ways. For example, children that go to school may be seen as investment goods for the household and therefore the children are better fed and cared for especially if more resources are available for the purpose. Going to school may also reduce in the risk factors that often increase the mortality of children through exposure to hazardous work conditions. All these new conditions would increase the survival of the children in the household. A more direct route is that of knowledge: children become educated and learn nutritional and hygiene that are put into practice either by themselves or by their parents that have now learned indirectly from their children. Once this knowledge is part of the household's human capital, and even in the absence of school in periods ahead, the human capital so far acquired in the household helps to maintain a better chance that children survival to adulthood.

$$
\begin{aligned}
&\tilde{n}_{it} = \tilde{n}_{it-1} \cdot \exp\left\{\kappa(s_j(\theta), \Lambda_{it})\right\} \\
&\textit{where} \\
&\qquad \Lambda_{it+1} = \Lambda_{it} \cdot \exp\left\{v(s_j(\theta), \Lambda_{it})\pi_{it}\right\} \cdot \exp\left\{\Delta \ln \tilde{n}_{it}\right\} \qquad if \quad e_{it} = 1 \\
&\qquad \Lambda_{it+1} = \Lambda_{it} \cdot \exp\left\{\Delta \ln \tilde{n}_{it}\right\} \qquad if \quad e_{it} = 0
\end{aligned} \tag{3}
$$

In expression (3) the parameter $v(s_j(\theta), \Lambda_{it})$ quantifies the household's human capital growth rate as result of the learning process due to the schooling attendance of children in the household. On the other hand the parameter $\kappa(s_j(\theta), \Lambda_{it})$ assumes that households of similar type have, on average, a similar rate of growth with regards to number and value of productive children thus assuming that the household's fertility rate and 'subjective value' of children is also conditional on levels of human capital.[19] The three dynamic expressions in (3) are assumptions and as such they can change according to different propositions regarding the evolution of the two endogenously defined variables.[20] With respect to the evolution of productive children ($\tilde{n}_{it}$), the expression is closely related to the classic assumption of population growth but it allows for such growth to vary heterogeneously between households that differ with regards to accumulation of human capital.[21] Expressions (2) and (3) provide the laws of motion that underline the endogenously changing state of the world.

3.4 The dynamic problem and the representative household

The fundamental problem faced by the representative household i is to choose optimal consumption and education so to maximize utility u_t subject to the endogenous evolution of assets (a_t), human capital (Λ_t) and number and value of children ($\tilde{n}_t$). The dynamic evolution is subject to stochastic and transitory productivity shocks π_t :

[19] Anticipating Sections 4.2-4.3 and Sections 5 and 6 is worth mentioning at this point that since children are the only contributors towards the growth of household's human capital, all the variability at household level that may affect fertility (e.g., parental age, existing number of children at levels, etc.) become determinants of human capital at levels.

[20] Appendix 2 proposes a variety of assumptions that may be equally valid at explaining the dynamic evolution of these variables and possible interactions among these.

[21] At a macroeconomic level the simplest form to express population growth is often $N_t = N_0 e^{\kappa t}$ where N_0 explains the initial population size and κ is the constant growth rate ($\dot{N}/N = \kappa$). We follow these structure but allow for heterogeneous human capital (Λ_{it}) – as opposed to time t – to be the factor that endogenously changes the size of productive children in the household. The assumptions implies $\dot{n}_t / n_t = \dot{\kappa}_{it}$ growth rate, that is, it is no longer assumed that households grow at some homogenous rate but instead growth is heterogeneous and subject to levels of accumulated human capital in the household.

$$
\begin{aligned}
&\max_{\{c,e\}_{t=0}^{T}} E_t \sum_{t=0}^{T}\left[\beta^t u_t(c_{it}) \middle| (X_{it},\psi_t)_{t=0}^{T}\right];\\
&where, \qquad X_{it} = \left(s_j(\theta), a_{it}, \Lambda_{it}, \tilde{n}_{it}, \pi_{it}\right)\\
&\qquad\qquad \psi_t = (r_t, W_{s_j(1)t},..., W_{s_j(\theta)t},..., W_{s_j(\bar{\theta})t}, \Upsilon_{s_j(1,n)t},...., \Upsilon_{s_j(\theta,n)t},..., \Upsilon_{s_j(\bar{\theta},n)t}, \tau_t)\\
&\qquad\qquad for \quad j=\{1,2,3,4\} \quad and \quad for \;\; each \;\; \theta \in \left\{1,2,..,\theta,...,\bar{\theta}\right\} \quad and \quad n \in \mathbb{N}^{+}\\
&and \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (4)\\
&a_{it+1} = (1+r_t)a_{it} + (e_{it})\left\{(1-\tau_t)\tilde{W}_{s_j(\theta)t}\pi_{it}\Lambda_{it}\right\} + (1-e_{it})\left\{(1-\tau_t)\left(\tilde{W}_{s_j(\theta)t}\pi_{it} + \tilde{\Upsilon}_{s_{ji}(\theta)t}\right)\Lambda_{it}\right\} - c_{it} \geq 0\\
&\qquad\qquad \tilde{n}_{it} = \tilde{n}_{it-1}\cdot \exp\left\{\kappa(s_j(\theta),\Lambda_{it})\right\}\\
&\qquad\qquad \Lambda_{it+1} = \Lambda_{it}\cdot\exp\left\{v(s_j(\theta),\Lambda_{it})\pi_{it}\right\}\cdot\exp\left\{\Delta\ln\tilde{n}_{it}\right\} \qquad\qquad if \quad e_{it}=1\\
&\qquad\qquad \Lambda_{it+1} = \Lambda_{it}\cdot\exp\left\{\Delta\ln\tilde{n}_{it}\right\} \qquad\qquad\qquad\qquad\qquad\qquad if \quad e_{it}=0
\end{aligned}
$$

Expression (4) describes a full dynamic structural model for optimal educational choices. Optimal allocation of lifetime resources implies maximizing expected discounted utility (i.e., $E_t\beta u_t(.)$) over the entire lifetime of the households that will be alive as long as children keep on growing towards adulthood and, in the process, accumulate human capital. Utility is obtained purely as result of current household consumption c_{it} that varies for the ith household according to number of household members. At each $t\in[0,T]$ over the finite horizon of the household the stochastic realization π_{it} determines the state of the world at t while future realizations remain undisclosed: only the distribution from which the shocks are drawn is assumed to be known and common to all households in the population. Once the realization is disclosed at t the household evaluates the alternative investment choice in the form of education and makes a mutually exclusive choice $e_{it}\in\{0,1\}$; the choice is optimal in that it takes into account the remaining expected lifetime utility implied by such choice and the effect that the choice may have on the evolution of the state of nature $X_{it+u,\,u>0}$ that itself emanates from the contemporaneous information resulting from the history of all past stochastic realizations up to the present time. Whereas $s_j(\theta)$ explains the household's start up human capital at the time of entry into productive activities ($t=0$), the household productive and investment history determines 'type specific' human capital Λ_{it} at any $t>0$. The state of the world is a function of both X_{it} that contains the household's heterogeneity and ψ_t that describes a vector of prices common (and common knowledge) to all households at t. Furthermore we define $X'_{it}=(a_{it},\Lambda_{it},n_{it})$ as the set of endogenous variables for the ith household of type $s_j(\theta)$ that result as consequence of optimal choices with regards to the only two control variables $\{c_{it},e_{it}\}$. Whereas consumption is assumed to be a continuous choice variable, the choice of education is modelled as a discrete choice made in each current period for as long as children keep on accumulating human capital.

The structural set up in (4) is a dynamic problem where stochastic shocks determine educational choices according to the combined effect that such innovations have on wealth and human capital. In the absence of shocks the model is a deterministic partial equilibrium model where households would remain in one of the two educational regimes (i.e., either $e_{it}=1$ or $e_{it}=0$) throughout the household's lifetime and according to start up human capital (i.e., household type and growth resulting from change in the number of productive children) that would determine initial choices at $t=0$ to be optimal for lifetime utility. Allowing for stochastic shocks explains an observed empirical regularity at the micro-economic level; agents make optimal choices at each point in time conditional on a changing world towards which the agent (the household) reacts idiosyncratically as result of its subjective preferences and characteristics. Thus, the household's evaluation of the innovations leads to potential shifts in educational regimes. Modelling the change in the state of nature as result of stochastic shocks can already be found in Kihlstrom and Laffont (1979) and Magnac and Rubin (1991, 1996); in these examples the realized state of the world leads to optimal regime choices in the labour market as result of attitudes towards labour market risks (see Section 2). This also implies that price uncertainty (i.e., uncertainty with respect to $W_{s(\theta)}$ and $\Upsilon_{s(\theta)}$) play no role in the process of selecting optimal household choices. Thus, our model is one where innovations reveal the contemporaneous state of nature (e.g., subjective believes on children's mortality, crop production opportunities, etc) that helps the household to exercise an educational choice (thus, a contemporaneous productivity choice) relative to the household's risk attitude that results from the household's own valuation of preferences, household characteristics and past educational history of the children in the household. Clearly, the choice to educate is a risky choice because contemporaneous decisions imply permanent investment effects in terms of human capital accumulation at the expense of current consumption. To illustrate the dynamics behind the structural model in (4) we assume a household characterized by $\Lambda_{is_3(low)t-1}$ at time $t-1$, i.e., an idiosyncratic house (i) where a single mother has completed some primary education ($s_3(low)$). The household has to make optimal choice at each point in time so to maximize current and expected lifetime utility conditional on all available (past and present) information. These choices include making an investment choice with regards to the schooling behaviour for the productive children in the household. Choices are only made once the household evaluates the state of the world that changes dynamically according unobserved stochastic shocks. We assume that the household receives a shock π_{it} at t so that once the innovation is evaluated the new state of the world suggests $e_{it}=1$ as the optimal choice conditional on the household's preferences and characteristics. That is, observing $e_{it}=1$ implies that relative to the household's preferences and characteristics the innovation π_{it} is evaluated as being 'high'. The choice implies that the combined receipts from productivity and human capital, i.e. $\left\{\left(W_{s_3(low)t}\pi_{it}\Lambda_{it}\right),\left(\Lambda_{it}\cdot\exp\left\{v(s_3(low),\Lambda_{it})\pi_{it}\right\}\exp\left\{\Delta\ln n_{it-1}\right\}\right)\right\}$, are viewed as higher relative to those implied by the alternative schooling regime $e_{it}=0$.[22] The choice implies that physical assets (a_{it}) and human capital ($\Lambda_{is_3(low)t}$) accumulate with the latter growing at rate $v(s_3(low), \Lambda_{is_3(low)t})$ and with clear permanent effects in terms of increased future potential productive capacity for the household. The human capital growth parameter $v(\cdot)$ depends on Λ_{it} because the learning process depends on actual human capital thus far accumulated in the household for a given household type. But if today's human capital increases as result of $e_{it}=1$ two things happen: children's survival becomes more likely and more important – thus the dependence of $\Delta\ln\tilde{n}_{it}$ on actual human capital in expression (4) – and the value of future investments on education increase alongside the increased value of having these children in the household. Therefore, once a schooling choice is made and human capital accumulates as result of such an investment, education in the future becomes more likely because the opportunity cost of no-schooling decreases relative to that of schooling.[23] A contrasting analysis can be made in the event of an adverse shock modelled as a very low value of the innovation π_{it}. For example, a bad harvest that drastically reduced the productive capacity of current endowments or some health shock that directly reduces adult's capacity to participate in the process of production: both examples could be modelled as $\pi_{it}\simeq 0$. Under such conditions the structural model in (4) implies that $e_{it}=1$ is not optimal because this leads to total returns $\left\{\left(W_{s_3(low)t}\pi_{it}\Lambda_{it}\right),\left(\Lambda_{it}\cdot\exp\left\{v(s_3(low),\Lambda_{it})\pi_{it}\right\}\exp\left\{\Delta\ln n_{it-1}\right\}\right)\right\}\Big|_{\pi_{it}=0}\simeq\left\{(0),\left(\Lambda_{it}\cdot\exp\left\{\Delta\ln n_{it-1}\right\}\right)\right\}$. Clearly, if the state of the world is determined by $\pi_{it}\simeq 0$ the optimal choice is that of no-schooling ($e_{it}=0$) because total receipts $\left\{\left(W_{s_3(low)t}\pi_{it}+\Upsilon_{s_3(low)t}\right)\Lambda_{it},\left(\Lambda_{it}\cdot\exp\left\{\Delta\ln n_{it-1}\right\}\right)\right\}$ are higher than those obtained from the alternative schooling choice $e_{it}=1$ that would now constitute a higher risk for two reasons: first, the adverse shock lowers the contemporaneous household production from adults thus increasing the value from children's complementary productive contribution at any given level of human capital. Secondly, a reduction in current production lowers the amount of resource that could be employed to enhance the survival of children; this means that expected gains from current investment in education are lowered as the value of the alternative choice in the form of children's productive output increase. But a current choice $e_{it}=0$ has permanent effect on the accumulation of future human capital because a reduction to both household's knowledge and children's survivals, thus making investing on education a riskier choice in periods ahead for any given $\pi_{it+u,u\geq 1}$ draw.

---

[22] Fro example, the shock might be a subsidy for each child that is newly sent to school, a sudden windfall in the form of a very good crop, an inheritance, etc.

[23] An educated child commands higher earnings because she is now more marketable. We notice that the price paid for the child (Υ_{it}) depends on n_{it} and not $\tilde{n}_{it}$ so that the price is now too low relative to the new productive capacity that having gone to school implies. In other words, once the child starts attending school, this lowers the

3.5 The Bellman representation

Expression (4) describes a dynamic problem fully determined by the existence of stochastic innovations π_{it} and the set endogenously changing state variables $(a_{it}, \tilde{n}_{it}, \Lambda_{it})$. The solution to this dynamic problem is a sequence of consumption choices $\{c_{it}\}_{t=0}^{T}$ each of which is associated with an admissible sequence of discrete educational regimes $\{e_{it} \in \{0,1\}\}_{t=0}^{T}$. These consumption choices are defined conditional on admissible initial and final conditions on values for the variables in $X_{it}' = (a_{it}, \Lambda_{it}, \tilde{n}_{it})$ that help to pin down admissible consumption paths (see Assumption 4 below). In order to define the sequences of consumption associated with the unique solution to (4) we choose to characterize the problem using recursive methods in terms of a value function $V_{it}^{s_j(\theta)}(\cdot)$. The structural set up in (4) assumes time independent shocks with the permanent effect of these picked up by physical and human capital, the only two components that carry contemporaneous information (that includes past experiences) to the future; it is important to emphasise that human capital (Λ_{it}) embeds the endogenous effects from $\tilde{n}_{it}$, that is, the growing number of children (fertility) interacted by the growing value of children as human capital accumulates through them. Thus, as functions of all three endogenous variables the structure in (4) provides the classic set up for the Bellman representation that relates a current value function $V_t(\cdot)$ to expectations of a future value function $V_{t+1}(\cdot)$ conditional on knowing all innovations up to t – i.e. $\{\pi_0, \pi_1, ..., \pi_{t-1}, \pi_t\}$ – where $V_{t+1}(\cdot)$ is discounted backward to time t assuming a discount factor β. Whereas $V_t(\cdot)$ explains the value of the maximised problem conditional on all paths at t for a given history of shocks, the Bellman representation introduces expectations on $V_{t+1}(\cdot)$ thus allowing for all expectation of future innovations $\{\pi_{t+1}, \pi_{t+2}, ..., \pi_{T-1}, \pi_T\}$. In view of that, the following Bellman representation characterizes the problem in (4) in the following set up:

$$V_{it}^{s_j(\theta)}(a_{it}, \Lambda_{it}, \tilde{n}_{it}, \pi_{it}, \psi_t) = \max_{\{c,e\}_{t=0}^{T}} \left\{ u_t(c_{it}) + \beta \cdot E_{\pi(t)}\left[V_{it+1}^{s_j(\theta)}(a_{it+1}, \Lambda_{it+1}, \tilde{n}_{it+1}, \pi_{it+1}, \psi_{t+1}) \right] \right\} \tag{5}$$

The value function in (5) summarizes the problem specific to each ‘type-and-ability’ household represented in the population. Thus, conditional on a given sub-population $s_j(\theta)$ the expression in (5) summarizes the current and future value obtained from choosing an optimal consumption path. The evolution of any given optimal path is subject to stochastic changes over the finite planning horizon: at each point in time households use all available information to form expectation on future realizations of

productive returns that children receive from non-schooling activities and, therefore, attending schools becomes ‘cheaper’: the cost of schooling is therefore lower and choosing school in future periods becomes more likely.

the innovation π thus updating the information set accordingly. Given discrete educational choices, there exists a unique solution characterizing the household's optimal consumption choice for each educational regime sequence $\left\{e_{it} \in \{0,1\}\right\}_{t=0}^{T}$ if and only if the value function in (5) is well behaved, that is, if the expression in (5) complies with a set of regularity conditions. These regularity conditions are subject to a set of assumptions necessary to proof that the value function in (5) is well behaved. We have already mentioned some of these conditions but we proceed to summarize these and other as follows:[24]

Assumptions

Assumption 1 (uncertainty): Stochastic productivity (also interpreted as educational) shocks π are assumed to be iid across time and households with a known, continuous and (at least once) differentiable distribution on a bounded non-negative support $[\underline{\pi}, \overline{\pi}]$.

Assumption 2 (utility): The homogeneous (across households) time varying utility function $u_t(\cdot)$ depends on consumption alone and changes heterogeneously among households as result of different consumptions needs. Thus we define $u_{it} = u_t(c_{it})$ and assume it to be strictly increasing, twice differentiable and concave in its single argument.

Assumption 3 (state space): The two state spaces spanned by the continuous state variables $X_{it} = (a_{it}, \Lambda_{it}, \tilde{n}_{it}, \pi_{it})$ and $X'_{it} = (a_{it}, \Lambda_{it}, \tilde{n}_{it})$ are assumed to be continuous, bounded and convex state spaces on a bounded support given by $[\underline{\pi}, \overline{\pi}]$. Household type $s_j(\theta)$ is part of the state space but assumed to be constant and exogenous throughout the planning horizon.

Assumption 4 (path identification): Initially, $a_{i0} = 0,\ \Lambda_{i0} = \underline{h}^{s_j(\theta)} > 0$ and $\tilde{n}_{i0} = n_0^{s_j(\theta)} \geq 1$. Terminal conditions implies $a_{iT} \geq 0,\ \Lambda_{iT} \geq \underline{h}^{s_j(\theta)} > 0$ and $\tilde{n}_{iT} > 0$. All three endogenous variables are assumed to take real values $\mathbb{R}$ that may include the zero only in the case of terminal physical assets a_{iT}.

Assumption 5 (non-crossing): The value function is assumed to have a derivative in the neighbourhood of zero that tends towards $-\infty$ from the right.

Assumption 6 (absolute risk aversion): Households display decreasing absolute risk aversion with risk attitudes towards educational choices that change inversely with changing assets and at a low pace. Let π^R represent the reservation level that determine the discrete

[24] The uniqueness of the solution in (5) depends on a set of regularity conditions leading to the description of a maximization problem where a concave function evolves over time subject to a set of monotonically changing constrains. The mathematical problem is similar to that in Lechner and Vazquez-Alvarez (2007); we draw from the latter to provide a modified set of assumptions to interpret the necessary conditions in the present paper.

educational choice so that π^R embodies the risk attitude towards education. Then, for any level of capital assets (a) decreasing absolute risk aversion implies that $(\partial\pi^R/\partial a) \leq 0$ while low pace change implies that $(\partial\pi^R/\partial a) \simeq 0$ in magnitude (but never exactly equal to zero).

Assumption 7 (growth and depreciation): The parameters $v(s,\Lambda)$ and $\kappa(s,\Lambda)$ are both constant for a given $s_j(\theta)$ type but change monotonically as result of changes on human capital. Human capital growth rates from educational choices are such that $v(s,\Lambda) \geq 0$: the parameter is assumed to increase monotonically with increasing levels of human capital Λ_{it} but at a decreasing rate. The parameter $\kappa(s,\Lambda) \geq 0$ can be positive value and it also changes monotonically according to changing levels of human capital.

Assumption 8 (Prices): All prices (i.e., W and Υ) are positive at any time period.

Assumption 9 (uniqueness): Consumption and savings are normal goods.

All assumptions (from 1 to 9) are required to determine a unique solution to (5) – and therefore to the problem in (4) – for any given educational regime choice and household type. It is therefore important to analyse each of these assumptions according to the behavioural meaning of the structural set up in (4). At any $t \in (0,T]$ nature randomly draws an innovation π_{it} that reveals the state of the world to the representative household of type $s_j(\theta)$ (Assumption 1). The risk avers household takes its preferences and characteristics (Assumption 6) as background information in order to evaluate the gains from choosing among alternatives regimes now allowing for a new state of the world that is well defined (bounded and convex, Assumption 3) and changes monotonically (Assumption 7). Choices with regards to consumption and education are made by rational households that maximize a well defined and known objective (Assumption 2) conditional on a monotonically changing dynamic state space (again, Assumption 3). The exclusion of leisure from the utility function (again, Assumption 2) abolishes wealth effects (i.e., backward bending labour supply functions) and this is consistent with populations where labour is inelastically supplied and substitution effects play no role in supplying labour at the household level. Allowing for lifetime constrains for all endogenous variables is a classic assumption in life-cycle models similar to the transversality condition in models of growth: these constrains (Assumption 4) identify feasible sets of consumption paths from which to choose the optimal one. These constrains are such that all household are homogeneous on physical assets at the start of their lives ($a_{i0} = 0$) but cannot die in debt at T (i.e., $a_{iT} \geq 0$) even if they are able to borrow over their existence (i.e., $a_{it} \geq 0$). Households with identical type $s_j(\theta)$ start their lives at $t = 0$ with an identical capacity to produce thus justifying the common endowment ($\Lambda_{i0} = \underline{h}^{s_j(\theta)}$) and the common evaluation of children ($\tilde{n}_{i0} = n_0^{s_j(\theta)}$) among households with an identical $s_j(\theta)$. The assumption implies an equal capacity to produce consumption goods among households with equal number of adults $s_i = s_j \in \{s_1, s_2, s_3, s_4\}$, an equal number of children at $t = 0$ (i.e., n_{i0}) and of equal ability type $\theta_i \in \{\theta_1, \theta_2, ..., \overline{\theta}\}$. All households are always capable to produce something ($\underline{h}^{s_j(\theta)} > 0$) and at the end of the household's life, at worst, the household still has the same capacity to produce as it had at $t = 0$ (basic learned knowledge for productive means is never lost so that $\Lambda_{iT} \geq \underline{h}^{s_j(\theta)} > 0$ applies). The underscore in $\Lambda_{i0} = \underline{h}^{s_j(\theta)}$ point out that at the beginning of the household's life human capital is at its lowest. As human capital changes monotonically (increasing or not) over the life-cycle, the concave value function may also change monotonically but may be further affected by exogenous shifts that could in principle lead to changes in the origin of this function: the situation is consistent with value functions for alternative educational regimes that may cross over time. To avoid this situation and, therefore, guarantee a unique solution we assume that all value functions originate in the neighbourhood of the zero point (Assumption 5). Together, Assumptions 2 and 5 emulate the Inada conditions on the value function. Finally, positive prices and positive human capital appreciation (Assumptions 7 and 8) are consistent with observed economic growth at the macroeconomic level and are necessary for the constrains to change monotonically over time. Assumption 9 is further required to characterize the Euler's condition in the presence of multiple interacting endogenous variables in the state space (see Section 3.6 and Appendix 1).

3.6 Properties of the structural model

The main focus on this paper is to identify the effect of education on growth (human capital growth rates) for distinct households that have been exogenously formed in conditions of adverse health environments. But identification and estimation of the main growth parameters in (4) rely on the consistency of such structural model. At the same time, evaluating the consistency of the structural set up is required in order to shape the identifying conditions on the parameters that we want to estimate. In what follows we provide a set of premises that define the properties of a unique solution to (5) and further help to provide identifying conditions for the parameters $v(\cdot)$ and $\kappa(\cdot)$. These premises require a more precise definition of π^R, a variable that was previously introduced as a reservation policy summarizing the preferences and characteristics of the household and determinant of regime choice (see Assumption 6). Clearly, households may be identical with regards to $s_j(\theta)$ but idiosyncratic preferences and characteristics act as conditionals, that is, identical $s_j(\theta)$ households react differently to a given innovation conditional on the actual state of the world. For example, two households (i and l) may be identical in type $s_j(\theta)$ but hold different information sets with regards to health knowledge (e.g., as result of the heterogeneous effects of HIV/AIDS information campaigns). The different information sets held by households i and l determine different subjective probabilities on their children's survival into adulthood and therefore (potentially) different schooling choices. At the same time, the information sets arise conditional on the state of the world that changes as consequence of the endogenously changing variables in the household (e.g., household's health knowledge may be a function of the household's wealth – that may determine communication access to information campaigns–, human capital or even the number of children in the household). Let π^R summarize all the household's preferences and characteristics that may be crucial at deciding the children's educational regime in the household and let's assume that π^R can change over time so that π_t^R applies. We have defined a situation where preferences between two households (e.g., i and l) differ, therefore we define π_{it}^R distinct from π_{lt}^R, or, in general, we let π_{it}^R be idiosyncratic. We have suggested that preferences and characteristics may change in some *deterministic* fashion according to a changed state of the world at t: this leads to a definition of π_{it}^R conditional on the known state space (X_{it}', ψ_t^T) where $X_{it}' = (a_{it}, \Lambda_{it}, \tilde{n}_{it})$ excludes the stochastic innovation and ψ_t^T explains the full price path from t to T, that is, $\pi_{it}^R\left(X_{it}^{'}, \psi_t^T\right)$.[25] Because $\pi_{it}^R\left(X_{it}^{'}, \psi_t^T\right)$ implies a summary of the household's preferences and characteristics it also implies a measure of the household's reaction toward a changing state of the world, that is, $\pi_{it}^R\left(X_{it}^{'}, \psi_t^T\right)$ is a 'benchmark' or *reservation level* (or *reservation policy*) that summarizes the household's fixed cost of educating the children.[26] We are now ready to introduce the required premises that determine the consistency of the structural set up. We start with Lemma 1 that characterizes the educational decisions for heterogeneous households.

Lemma 1 (regime choice): Assumptions 1-8 hold. Let π_{it} be an innovation that determines the state of the world at $t \in (0,T]$. Conditional on π_{it}, an optimal educational choice is characterized by a monotonic educational reservation policy that is unique to each household at time t according the household preferences and characteristics such that,

[25] We have defined expressions (4) and (5) so that once the π_{it} realization is known the state of the world $X_{it}^{'}$ is fully defined. Allowing for π^R to be further conditioned on the stochastic shock would imply the assumptions that household's preferences are stochastic (vary with the throw of a dice) which is not consistent with rational human behaviour. Therefore it is more consistent with human behaviour to let π^R be a function of $X_{it}^{'}$ as opposed to X_{it}. Assuming uncertainty over price paths implies that agents know the sequence ψ_t^T.

[26] Anticipating Sections 4 and 5 it can be said that $\pi_{it}^R\left(X_{it}^{'}, \psi_t^T\right)$ may be identified using all observed characteristics in the household that vary over time. An analogy may be made between the classic reservation wage that determines the unemployed individual fixed cost of entry in the labour market and $\pi_{it}^R\left(X_{it}^{'}, \psi_t^T\right)$ that determines the household's fixed cost of entry into the education system.

$$\begin{aligned}&\textit{The household prefers } e_{it} = 1 \textit{ if}\\&\qquad \pi_{it} > \pi^{R}_{it}\left(X^{'}_{it}, \psi^{T}_{t}\right) \qquad (6)\\&\textit{Alternatively, the household prefers } e_{it} = 0.\end{aligned}$$

Proof, **see Appendix 1**

Lemma 1 and its proof in Appendix 1 underline a tractable value function that remains flexible enough to allow for multiple solutions (i.e., multiple consumption paths consistent with the structural set up in expression (4)). In fact, Lemma 1 characterizes the educational decision such that the value function in (5) can be interpreted as the sum of mutually exclusive value functions each associated with a unique educational regime over the finite support $[\underline{\pi}, \overline{\pi}]$.[27] Lemma 2 below constrains the problem to a set of value functions consistent with rational households that derive utility from increasing levels of consumption at a decreasing rate (i.e., the value function is concave in assets):

Lemma 2 (concave value function): Assumptions 1-8 hold. The expected value function $E^{s_j(\theta)}_{\pi} V_{t+1}(a, \Lambda, \tilde{n}, \pi)$ is a strictly increasing, continuous (twice differentiable) and concave function of assets (a_{t+1}) for any given level of (continuously changing) human capital and number of children.

Proof, **see Appendix 1**

The Bellman representation in (5) shows that the only change to the state of the world between time periods is due to the stochastic shock π as emphasised by indexing the expectation as E_{π}. The stochastic change determines the new state of the world (including possible changes to the preferences and characteristics in the household) and, therefore, the possibility to change the educational regime e_{it}. However, the household obtains utility not from the educational choice but directly from the choice of some consumption bundle c_{it} at each point in time. The educational choice simply conditions the optimal path so that the solution to the problem in (5) is an optimal consumption path *for each* of the education regimes considered (i.e., for $e_{it} = 0$ and $e_{it} = 1$).[28] Following from the classic Bellman representation set

[27] We notice that the value function varies stochastically over the support of π as result of its dependence on the expectations of such stochastic shocks. Therefore, the value function is defined of the support of π that we have defined as $[\underline{\pi}, \overline{\pi}]$ (see Assumption 1).

[28] If we allow for the educational choice to be some continuous variable as opposed to a discrete choice, the solution to the problem would be a continuum of consumption paths over the continuous values of the educational choice variable (e_{it}). Instead of a single path over time the solution is a continuous surface of optimal solutions. Still, utility would come from consumption alone because this is the only argument in the utility function. Clearly, letting e_{it} be a continuous variable in the interval $\{0,1\}$ makes the problem more complex but implies a solution that does not provide a qualitative improvement in terms of insight.

up, each of these optimal paths is characterized by an Euler Equation that explains the intertemporal consumption rule conditional on educational choices. Characterizing the solution with an Euler Equation implies that such equation fulfils the set of regularity conditions described in Lemmas 1 and 2. These regularity conditions are sufficient and necessary so that the Euler specification defines optimal consumption decisions for fixed educational choices, with the conditional Euler Equation expressed as follows:

$$\left.\frac{\partial u_t(c_{it})}{\partial c_{it}}\right|_{e_{it}} = E_\pi\left[\left\{\beta(1+r_t)\right\}\cdot\left.\frac{\partial u_{t+1}(c_{it+1})}{\partial c_{it+1}}\right|_{e_{it}}\right] \tag{7}$$

The solution to (5) as characterized by (7) does not account for the fact that other (endogenous) variables change alongside assets (a). These variables are human capital (Λ) that at the same time depends on endogenously changing number and value of children ($\tilde{n}$). We assume that children are potentially productive units. Thus, it is not so much 'how many productive children' does the household have but how much productivity a child can bring to the household. Productivity is based on both the number of children and the productive level of children defined according to human capital: this further reinforces the fact that the growth rate of children in expression (4) is given as a continuous variable rather than a step function. Thus, for a given stochastic shock, continuous assets move along continuous changes on human capital and number of productive children. Unless we account for these changes on the state space, the characterization of the consumption rule in (7) is not sufficient for a unique solution. Lemma 3 together with its proof in Appendix 1 are further required so that the Euler Equation in (7) implies a unique solution to the optimal path conditional on given levels of human capital with this latter dependent on the endogenously changing number of children:

Lemma 3 (uniqueness of the optimal consumption path): Assumptions 1-9 hold. The Euler Equation in (7) is not sufficient to identify an optimal consumption path in the presence of endogenous changes in human capital and number of productive children. Identification further requires that consumption and savings are normal goods for fixed educational decisions.

Proof, **see Appendix 1**

4. Identification and Estimation

This section draws from the structural model in Section 3 to present the conditions required to identify the parameters $v(\cdot)$ and $\kappa(\cdot)$ allowing for constrains implied by a unique solution to the representation in (5). Once the identification conditions for the key parameters have been established we define a flexible nonparametric based estimation procedure that embodies the identification strategy to estimate the key parameters regarding the accumulation of human capital. Identification and estimation relies both on the structural set up from Section 3 and on assumption that the data is sufficiently informative so that it can help identify the human capital parameters for given populations. We start by defining a set of reduced form specifications that provide an explicit interpretation in terms of human capital growth rates.

4.1 Educational choices and reduced form specifications

A set of reduced form specifications can be directly drawn from the structural set up in (4). Each reduced form provides an explicit interpretation of the contemporaneous production in the household as result of the household's behaviour regarding an educational choice. Our modelling strategy assumes that education is a binary choice so that there are two possible reduced forms specifications; one associated with the contemporaneous productivity outcome in the event that all productive children in the household are sent to school, and an alternative reduced form specification resulting from a contemporaneous regime choice with a non-educational status for the productive children in the household. Each of the following sub-sections describes each of the two reduced form specifications and in each case provides the interpretation of these reduced forms in terms of human capital growth rate.

Regime 1: Education

From the structural set up in (4), a household that chooses to invest on education at t (i.e., $e_{it} = 1$) obtains total *net earnings* equal to $E_{it}^{s_j(\theta)} = (1-\tau_t)\tilde{W}_{s_j(\theta)t}\pi_{it}\Lambda_{it}$, simplified to $E_{it}^{s_j(\theta)} = W_{s_j(\theta)t}\pi_{it}\Lambda_{it}$, where the suffix (it) and the indexing $s_j(\theta)$ imply that earning are heterogeneous to the ith household at t but otherwise common to household's of identical class with regards to type s_j and skill specific ability θ. Given that $(i, s_j(\theta))$ remains constant over time for the ith representative household, from now on we drop these indicators to simplify the notation throughout Section 4.[29]

Net earnings E_t amount to a household-class and skill-specific net average return W_t for each unit of contemporaneous human capital Λ_t and augmented by the time-varying productivity shock π_t that we modelled as having a direct effect on the productive capacity of decision makers in the household but only affects build up knowledge in the event of a schooling choice (see Section 3.4). Therefore, alongside net earnings E_t – and in contrast with a non-schooling choice –, when children in the household go to

[29] With this all expressions that follow refer a representative household of any given type as defined by number of adults and average skill class.

school ($e_{it}=1$) this enhances human capital at the rate of $v(\Lambda_t)$, i.e., $\Delta\ln\Lambda_{t,t+1}=v(\Lambda_t)\pi_t+\kappa(\Lambda_t)$; the value $\kappa(\Lambda_t)$ is the rate at which the household's human capital is enhanced (decreased) as result of augmenting (decreasing) n_t even in the absence of schooling and resulting from both fertility outcomes and human capital in the household.[30] Clearly, observed household net earnings at levels (E_t) as result of a contemporaneous schooling choice are not informative about human capital growth rates ($\Delta\ln\Lambda_{t,t+1}$). However, earnings' growth that are observed over a spell of *at least* two consecutive educational periods (i.e., households such that $e_t=e_{t+1}=1$) are informative about $\Delta\ln\Lambda_{t,t+1}$ as shown in the following expression:

$$\begin{aligned}
&E_t = W_t\pi_t\Lambda_t \qquad \textit{where} \qquad && \Lambda_{t+1}=\Lambda_t\cdot\exp\{v(\Lambda_t)\pi_t\}\cdot\exp\{\Delta\ln\tilde{n}_t\}\\
& && \tilde{n}_t=\tilde{n}_{t-1}\cdot\exp\{\kappa(\Lambda_t)\}\\
& && E_{t+1}=W_{t+1}\pi_{t+1}\Lambda_{t+1}\\
&\Rightarrow && \Delta\ln E_{t,t+1}=\Delta\ln W_{t,t+1}+\Delta\ln\pi_{t,t+1}+\left[\ln\Lambda_{t+1}\left(\pi(t)\right)-\ln\Lambda_t\left(\pi(t-1)\right)\right]\\
&\Rightarrow && \Delta\ln E_{t,t+1}=\Delta\ln W_{t,t+1}+\Delta\ln\pi_{t,t+1}+\Delta\ln\Lambda_{t,t+1}\\
&\Rightarrow && \Delta\ln E_{t,t+1}=\Delta\ln W_{t,t+1}+\Delta\ln\pi_{t,t+1}+\left(v(\Lambda_t)\pi_t+\kappa(\Lambda_t)\right)
\end{aligned} \tag{8}$$

Expression (8) draws from (4) to show the consequence on the growth of earnings when a household chooses education consecutively over two periods. The last line in (8) shows that three distinct components contribute towards earnings' growth: (i) growth in wages ($\Delta\ln W_{t,t+1}$) specific to a given type of households,[31] (ii) growth in earnings as consequence of idiosyncratic changes in human capital ($\Delta\ln\Lambda_{t,t+1}$) and (iii) growth in earnings as consequence of stochastic productivity gains as direct consequence of choosing a spell of education over two consecutive periods. Expression (8) explicitly shows that Λ_{t+u} can also be expressed as $\Lambda_{t+u}\left(\pi(t+u-1)\right)$ for $u=0,1$. The term $\pi(t+u-1)$ should not be confused with π_{t+u-1} (i.e., the variable π at time $t+u-1$), but instead we should read $\pi(t+u-1)$ as the summarized effect of productivity and educational history of the household up to time $t+u-1$. The key point to make here is that all educational choices from $t=0$ onwards result from the household's subjective evaluation of a sequence of time-independent stochastic shocks so that shocks up to time t are

[30] Our model strategy implies a continuum number of children as result of their value as productive units with productivity increasing in a concave fashion as result of concave human capital.

[31] Although we suppress the index $s_j(\theta)$ throughout, the variables W_t and W_{t+1} should be read as specific to the class of households defined by $s_j(\theta)$ such that W_t emulates $W_t^{s_j(\theta)}$ and likewise for W_{t+1}. Wages are specific to household class but the variables π_t, π_{t+1}, $\tilde{n}_t$ and $\tilde{n}_{t+1}$, as well as the parameters v_t and κ_t, are all class specific ($s_j(\theta)$) and furthermore specific to each household (i).

responsible for shaping Λ at time $t+1$: expressing Λ_{t+u} in its longer version $\Lambda_{t+u}\left(\pi(t+u-1)\right)$ – for $u=0,1$ – puts emphasis on the fact that human capital accumulates as consequence of the household's history (up to the previous period) of productive and educational choices that results from a subjective evaluation of an stochastically changing state of the world. Lemma 1 implies that a household's subjective evaluation is associated with a reservation policy π^R resulting from the household's preferences and characteristics. Thus, Λ at time $t+1$ depends on the history of stochastic shocks up to time t where such shocks are evaluated against a history of time-dependent reservation policy values π_t^R also up to time t. Whereas human capital Λ and stochastic shocks π are unobserved to the econometrician, variables that may explain household preferences and characteristics (and, therefore π_t^R at any t) are observed. Thus, when Λ_{t+u} is interpreted as $\Lambda_{t+u}\left(\pi(t+u-1)\right)$ in (8), we are emphasising a link between observed variability (π^R) and unobservable variables (π, Λ), a link that becomes crucial at the point of identifying human capital growth rates in sections ahead. The last line in (8) explicitly shows the link provided by (4) between growth rates in $\Lambda_{t+u}\left(\pi(t+u-1)\right)$ and the parameters $v(\Lambda_t)$ and $\kappa(\Lambda_t)$: the two parameters are the central pieces in our research in the sense that having theoretically justified the structural model our empirical aim would be to estimate these two parameters for each household class (type and skill specific, $s_j(\theta)$) represented in given populations of interest. The parameter $v(\Lambda_t)$ explains growth in the productive capacity of the household as result of the household specific 'ability to learn' that results from an optimal schooling choice. The parameter $\kappa(\Lambda_t)$ complements human capital growth not as result of learning but as result of the change in the number of productive children in the household and their quality as present and future productive units as described by the fact that $\kappa(\Lambda_t)$ is explained by Λ_t. Both parameters $v(\Lambda_t)$ and $\kappa(\Lambda_t)$ vary over time as consequence of their dependence on realized contemporaneous human capital *at levels*. Clearly, human capital growth rate from learning ($v(\Lambda_t)$) depends on the level of human capital in the household because 'by how much' will human capital grow tomorrow depends on how much knowledge there is the household today. This follows directly from the classic assumption in labour economics that human capital is concave over time. The fact that $\kappa(\Lambda_t)$ depends on human capital at levels implies that 'productive fertility' in the household is also concave: again, this further illustrates that we are not equating the number of productive children at t (i.e., n_t) to the actual discrete number of children; instead, the actual value of the children imply additional productive units in the household that may monotonically increase (or decrease) in some concave fashion alongside human capital at levels.

Regime 2: No Education

When the optimal choice for the household is that of no education at t (i.e., $e_t = 0$), net earnings at levels (Γ_t) are given by $\Gamma_t = \left(W_t\pi_t + \Upsilon_t\right)\Lambda_t$ – see expression (4). As was the case with the previous regime – Regime 1 –, W_t explains the household class and skill-specific gain from each unit of productive (human) capital Λ_t in the household, with household specific skill prices W_t augmented by the stochastic productivity shock π_t of decision makers in the household. Expression $\Gamma_t = \left(W_t\pi_t + \Upsilon_t\right)\Lambda_t$ implies that the only reason why children are not sent to school is because they are needed to meet the consumption needs in the household. When parents take their productive children to work – or likewise send all productive children to work elsewhere – the earnings in the household increase by a factor Υ_t per unit of productive (human) capital Λ_t. Unlike W_t, the factor Υ_t is household specific as it depends on the number n_t of productive children at t, a number that can vary over time and is therefore specific to each household irrespective of $s_j(\theta)$. The effect of the augmenting factor π_t on Υ_t is missing because productive shocks directly affect both the productivity of household's decision makers and the choices they make for the household. This implies that children's contribution (Υ_t) act as some safety net in the event of adverse shocks to the productive capacity of adults and(or) endowments in the household, an assumption that helps us make an analogy between Υ_t in developing economies and unemployment benefits for the under-employed in the developed world where labour market insurances are well defined. However, an adverse innovation π_t does affect children indirectly for two reasons: first, contemporaneous absence from formal schooling implies that children (and therefore the household as a unit) cannot experience the gains from the augmented productive capacity that results from stochastic changes that contemporaneously affects only those who attend school (i.e., from learning). Secondly, for any given level of Λ_t, human capital in the future (Λ_{t+1}) suffers as result of lost contemporaneous learning (i.e., $v(\Lambda_t)$) with current loss in learning feeding into human capital in periods ahead (e.g., Λ_{t+2}) through lagged effects given by $\kappa_{t+1}(\Lambda_{t+1})$. This effect at t is permanent and for periods ahead even if the household restores a schooling regime at $t+1$, a loss that could be interpreted as the adverse effect that lower levels of human capital has on children's survival into the future. Similar to the situation implied by a regime under education, earnings at levels for households with $e_t = 0$ are not sufficient to identify the human capital growth parameter $\kappa_t(\Lambda_t)$. However, observing households in spells such that $e_t = e_{t+1} = 0$ provide a reduced form specification that helps to isolate the parameters of interest:

$$\Gamma_t = (W_t\pi_t + \Upsilon_t)\Lambda_t \qquad \textit{where} \quad \Lambda_{t+1} = \Lambda_t \cdot \exp\{\Delta \ln \tilde{n}_t\}$$
$$\tilde{n}_t = \tilde{n}_{t-1} \cdot \exp\{\kappa(\Lambda_t)\}$$
$$\Gamma_{t+1} = (W_{t+1}\pi_{t+1} + \Upsilon_{t+1})\Lambda_{t+1} \tag{9}$$
$$\Rightarrow \qquad \Delta \ln \Gamma_{t,t+1} = \Delta \ln(W\pi + \Upsilon)_{t,t+1} + \left[\ln \Lambda_{t+1}(\pi(t)) - \ln \Lambda_t(\pi(t-1))\right]$$
$$\Rightarrow \qquad \Delta \ln \Gamma_{t,t+1} = \Delta \ln(W\pi + \Upsilon)_{t,t+1} + \Delta \ln \Lambda_{t,t+1}$$
$$\Rightarrow \qquad \Delta \ln \Gamma_{t,t+1} = \Delta \ln(W\pi + \Upsilon)_{t,t+1} + (\kappa(\Lambda_t))$$

As was the case with expression (8), expression (9) provides an explicit interpretation for the human capital parameter $\kappa(\Lambda_t)$. The interpretation of this parameter is identical to that under Regime 1 but may differ in magnitude under the distinct educational regime. The last equality in expression in (9) shows that between periods' growth in earnings for the households $(\Delta \ln \Gamma_{t+1})$ amounts to the sum of two components: (i) productivity augmented household's class-and-skill specific growth in W with additions from growth in Υ_t that comes from children's contribution towards production, and (ii) *the changing rate of human capital* as result of idiosyncratic changes to productive capacity because of changes in the number and quality of productive children. With respect to (i) the difficulty arises at the empirical level; we notice that in the case of Regime 1 the term $\Delta \ln W$ does not include idiosyncratic household changes so that it can be identified as the average change in the skill price (e.g. as measured by basic average wages) within the subset in the population that share the same $s_j(\theta)$ class. Adding items that are household specific as result of Υ_t's dependence on n_{it} implies that one cannot disentangle $W\pi$ from $(W\pi + \Upsilon)$ unless we define Υ more explicit as function of the pertinent household's characteristics. To do so, we let $\Upsilon_t = W_t\sqrt{n_t}\left(\eta(\bar{g}_t)\right)$ where the value $\eta_t = \eta(\bar{g}_t) \in (0,1)$ explains children's productivity with respect to that of adults and where we assume that η_t depends on the (time varying) average age of the children in the household at time t, i.e., $(\bar{g}_t)$.[32] For example, when children in the household are very young but of schooling age (e.g., $\bar{g}_t \simeq 5$) their contribution toward production relative to that of an adult will be $\eta_t \simeq 0$. As children age and become almost as productive as adults (e.g., $\bar{g}_t \simeq 15$), $\eta_t \simeq 1$: it is assumed that $\eta(\bar{g}_t)$ is identical for all households whose children belong to the same generation. Letting the number of children enter in terms of $\sqrt{n_t}$ is a way of modelling the fact that parents would be paid more for an increasing number of children in the workplace but at a decreasing rate as more children are

[32] The specification draws from a calibrating assumption fund in the GE-model of Bell, Devarajan and Gersbach (2006). In their case the assumption is simplified such that $\Upsilon_t = W_t n_t \eta$ with η common for all age groups since in a macroeconomic set up the average age of children in the household (and household characteristics in general) are lost in aggregation.

added in the production process.[33] At this point we notice that we have not used $\tilde{n}_t$ but instead we use the actual number of children n_t: this further reinforces the difference between number of children and their productive value in terms of household's investment.[34] Allowing for $\Upsilon_t = W_t \sqrt{n_t}\left(\eta\left(\bar{g}_t\right)\right)$ helps to characterize the household's idiosyncratic effect in Υ as that implied by n_{it} so that with complete indexing Υ would be given by $\Upsilon_{it}^{s_j(\theta)} = W_{s_j(\theta)t}\sqrt{n_{it}}\left(\eta\left(\bar{g}_t\right)\right)$ where $W_{s_j(\theta)t}$ remains as specific to household with identical type. The specification transforms the last expression in (9) as follows:

$$
\begin{array}{ll}
\textit{From } (9), & \Delta\ln\Gamma_{t,t+1} = \Delta\ln(W\pi + \Upsilon)_{t,t+1} + \left(\kappa(\Lambda_t)\right) \\
\textit{where} & \Upsilon_{t+u} = W_{t+u}\sqrt{n_{t+u}}\,\eta\left(\bar{g}_{t+u}\right), \quad \textit{for} \quad u = 0,1 \\
\textit{Simplify}; & \Upsilon_{t+u} = W_{t+u} n_{t+u}^{1/2}\eta_{\bar{g}(t+u)}, \quad \textit{for} \quad u = 0,1 \\
\Rightarrow & \Delta\ln\Gamma_{t,t+1} = \left\{\ln\left(W_{t+1}\left(\pi_{t+1} + n_{t+1}^{1/2}\eta_{\bar{g}(t+1)}\right)\right) - \ln\left(W_t\left(\pi_t + n_t^{1/2}\eta_{\bar{g}(t)}\right)\right)\right\} + \left(\kappa(\Lambda_t)\right) \\
\Rightarrow & \Delta\ln\Gamma_{t,t+1} = \Delta\ln W_{t,t+1} + \left\{\ln\left(\pi_{t+1} + n_{t+1}^{1/2}\eta_{\bar{g}(t+1)}\right) - \ln\left(\pi_t + n_t^{1/2}\eta_{\bar{g}(t)}\right)\right\} + \left(\kappa(\Lambda_t)\right) \\
\Rightarrow & \Delta\ln\Gamma_{t,t+1} = \Delta\ln W_{t,t+1} + \left\{\Delta\ln\left(\left(\pi + n^{1/2}\cdot\eta_{\bar{g}}\right)\right)_{t,t+1}\right\} + \left(\kappa(\Lambda_t)\right)
\end{array}
\tag{10}
$$

The last expression in (10) is such that $\Delta\ln\left(\left(\pi + n^{1/2}\cdot\eta_{\bar{g}}\right)\right)_{t,t+1}$ depends on the value of π_{t+1} relative to π_t; however, for low values of productivity associated with adverse innovations (e.g., if $\pi \simeq 0$), the term $\Delta\ln\left(\left(\pi + n^{1/2}\cdot\eta_{\bar{g}}\right)\right)_{t,t+1}$ may be positive if $n_{t+1} > n_t$ but negative if $n_{t+1} < n_t$ (childhood mortality) over short spells such that $\eta_{\bar{g}_{t+1}} \simeq \eta_{\bar{g}_t}$. Furthermore, in the event of adverse shocks that affect the household's productive capacity, when the number of children remains constant over consecutive time periods $(n_{t+1} = n_t)$ growth in earnings as result of children's contribution will be from the children's capacity to earn more (monetary gain) as they become older, i.e., from changes such that $\Delta\eta_{\bar{g}_{t,t+1}} > 0$. The last expression in (10) clearly distinguishes between children's contribution to earning's growth as result of the 'number of children' in the household for given productivity shocks, that is, $\Delta\ln\left(\left(\pi + n^{1/2}\cdot\eta_{\bar{g}}\right)\right)_{t,t+1}$, separated from the contribution that children imply as investment units capable to survive for periods ahead, that is, $\kappa(\Lambda_t)$. Indeed, in the case of no education, the value $\kappa(\Lambda_t)$ is defined as *the changing rate*

[33] One could make other less restricted assumptions so that instead we allow for $\sqrt[u]{n_t}, u \in [1,2)$ thus increasing the gains from more number of children and including the choice $\sqrt[u]{n_t} = n_t$.

[34] For example, a producer that employs both the parents and the children is only interested on the *number* of children (n_t) that contribute towards production today, and not on the potential human capital that children imply for the future.

of human capital if the number of children increases or if the human capital changes over time for the same number of children: the definition opens the possibility that $\kappa(\Lambda_t)$ may not increase as result childhood mortality or stagnated human capital over consecutive time periods for the same number of children. In the event that the number of productive children (strictly) decreases over time as result of mortality the rate of change for $\tilde{n}_{it}$ may not necessarily be negative but Υ_t would decreases over consecutive time periods. The identification conditions in Section 4.2 and 4.3 provide data driven assumptions to isolate and thus identify the parameters of inters, but the actual empirical values of these parameters depend on the empirical application.

4.2 Identification strategy

Expressions from (8) and (10) provide reduced form specifications that explicitly show the human capital parameters conditional on household type. Both expressions draw directly from the structural model in (4) while relying on observed earnings as best signal to identify the dynamic productivity choices that households make (see Section 3). However, as they are now defined, expressions in (8) and (9) are not yet operational at an empirical level. The reason for this is that human capital parameters $v(\Lambda_t)$ and $\kappa(\Lambda_t)$ are given conditional on Λ_t that is not directly observed from the data. Likewise, knowledge of the magnitude of the innovation set (π_t, π_{t+1}) is required to estimate the value $v(\Lambda_t)$ for any given household type when in fact the innovations π_t for any given t are defined in the model as unobserved variability. On the other hand $v(\cdot), \kappa(\cdot)$ and π_t all vary over time but are jointly defined in expression (8) so that identification also implies conditional assumptions that disentangle the effect between parameters and the innovations. The purpose of this section is to provide a set of valid data-driven assumptions that will lead to the identification of the key human capital formation parameters. We start by defining two population moment conditions one for each educational regime and conditional on (omitted index) household type:

$$\begin{aligned}
&\textit{From } (8): \quad \textit{Regime } 1 \quad (\textit{education}) \\
&\mathrm{E}\left[\Delta\left(\ln E_{t,t+1}\right) \middle| e_t = e_{t+1} = 1\right] = \\
&\mathrm{E}\left[\Delta\left(\ln W_{t,t+1}\right) \middle| e_t = e_{t+1} = 1\right] + \mathrm{E}\left[\Delta\left(\ln \pi_{t,t+1}\right) \middle| e_t = e_{t+1} = 1\right] + \mathrm{E}\left[\left(v(\Lambda_t)\pi_t + \kappa(\Lambda_t)\right) \middle| e_t = e_{t+1} = 1\right]
\end{aligned} \tag{11}$$

$$\begin{aligned}
&\textit{From } (10): \quad \textit{Regime } 2 \quad (\textit{no education}) \\
&\mathrm{E}\left[\Delta\left(\ln \Gamma_{t,t+1}\right) \middle| e_t = e_{t+1} = 0\right] = \\
&\mathrm{E}\left[\Delta\left(\ln W_{t,t+1}\right) \middle| e_t = e_{t+1} = 0\right] + \mathrm{E}\left[\Delta \ln\left(\pi + n^{1/2} \cdot \eta_{\bar{g}}\right)_{t,t+1} \middle| e_t = e_{t+1} = 0\right] + \mathrm{E}\left[\kappa(\Lambda_t) \middle| e_t = e_{t+1} = 0\right]
\end{aligned}$$

In expression (11) the expectations operator $\mathrm{E}[\cdot]$ should not be confused with the variable earnings growth (E_t) for households with educational spells. Whereas Section 4.1 shows that the population moment conditions in (11) are informative on the parameters underlying human capital formation, this section provides identifying conditions so that the sample analogue to the population conditions in (11) are use to obtain empirical estimates of the key parameters. Our first set of identification strategies deals with conditions that imply disentangling the contributions from $v(\Lambda_t)\pi_t$ and $\kappa(\Lambda_t)$; so far both are assumed to vary as result of the same variation on human capital over time. The second set of identifying assumptions provides conditions that disentangle the effects between $v(\Lambda_t)$ and π_t since the two terms are given multiplicatively and varying over time. Our third and final set of identification strategies deals with the elicitation of the two unknown variables (Λ_t and π_t).

We start by defining a set B_t (for any t) that includes determinants of the parameter $\kappa(\Lambda_t)$ for a given level of human capital Λ_t. We define the set B_t such that conditional on Λ_t children's value to the household is identical in both regimes (education or not). For example, two households (i,l) signal identical levels of $\Lambda_{it} = \Lambda_{lt} = \Lambda'$ but experience different productivity shocks such that we observe the choice sets to be ($e_{it} = e_{it+1} = 1$) and ($e_{lt} = e_{lt+1} = 0$) respectively. The moment conditions in (11) imply that subpopulation expectations originate from $E[v(\Lambda')\pi + \kappa(\Lambda') \mid e_{it} = e_{it+1} = 1]$ and $E[\kappa(\Lambda') \mid e_{lt} = e_{lt+1} = 0]$ for each respective household. For known π (assume this for now) nothing in our model suggest that the values $E[\kappa(\Lambda') \mid e_{it} = e_{it+1} = 1]$ and $E[\kappa(\Lambda') \mid e_{lt} = e_{lt+1} = 0]$ should be the same. However, in the absence of human capital our structural set up suggests that determinants of household's fertility would determine $\kappa(.)$. Thus, let B_t be a set of variables that contains specific information about the household's fertility position within the finite life-cycle in the household. For example, B_t may contain a set of dummy variables that explains the number of children at t, the cohort of children at t and variables that identify the age of parent and therefore their ability to reproduce. Conditional on B_t and for a similar level of human capital the value $\kappa(\Lambda')$ is identified for ($e_{it} = e_{it+1} = 1$) away from $v(\Lambda')$ because within cells defined by Λ' we may equate $E[\kappa(\Lambda') \mid e_{lt} = e_{lt+1} = 0] = E[\kappa(\Lambda') \mid e_{lt} = e_{lt+1} = 1]$. Assumption 10 summarizes this data-driven identifying condition:

Assumption 10 (homogeneity in fertility): Within each $s_j(\theta)$ class that signals an equal value of Λ'_t, a set $\mathrm{B}_t = \mathrm{B}_{t,E} = \mathrm{B}_{t,\Gamma}$ contains observed information defining sub-groups in the population with homogenous fertility histories. Conditional on membership to the set B_t (for given Λ'_t) human capital growth due to the evaluation of surviving children is identical between educational regimes so that $E[\kappa(\Lambda')\,|\,\mathrm{B}_t] = E[\kappa(\Lambda')\,|\,\mathrm{B}_t, e_{lt} = e_{lt+1} = 0] = E[\kappa(\Lambda')\,|\,\mathrm{B}_t, e_{lt} = e_{lt+1} = 1]$ applies.

Assumption 10 implies that once we have sub-divided the population of households with ($e_{lt} = e_{lt+1} = 0$) choices by $s_j(\theta)$ and obtain $\kappa(\Lambda)$ from the sample analogue to (11), we may be able to identify $v(\Lambda)\pi$ away from $\kappa(\Lambda)$ for individuals in ($e_{it} = e_{it+1} = 1$) with identical B_t identifiers within sub-divisions given by $s_j(\theta)$. Next, we have to deal with the problem of disentangling the effects between the two time varying terms $v(\Lambda_t)$ and π_t. For this purpose we also use a new data-driven assumption that will require the existence of informative data sets. Let Z_t be a set of variables that determines Λ_t. There are many candidates for this particular set: years of experience in the labour market, age of parents, presence and characteristics of grandparents in the household, type of labour market characterising the production in the household (e.g., agriculture, traders, etc) and, on behalf of children, the schooling history of children in the household. We notice that Z_t may also include B_t since fertility history in the household is part of the households human capital accumulation. In general, we assume that Z_t contains sufficient information so that within cell defined by Z_t households of identical type (i.e., by $s_j(\theta)$ sub-population) human capital is sufficiently homogenous so that within cells the household children are equally capable to learn as result of the schooling choice exercised by their parents. Thus, conditional on Z_t for a specific $s_j(\theta)$ household type, $v(\Lambda_t) = v_t$ remains constant for that particular time period t for all households in that data-driven cell even if each specific household receives a distinct innovation π_t. Assumption 11 summarizes the identifying condition as follows:

Assumption 11 (homogeneity in human capital): Within each $s_j(\theta)$ class the parameter $v(\Lambda_t)$ explains growth due to education under schooling choices. Let Z_t be a set of observed variability that explains sub-population with homogenous levels of human capital for specific $s_j(\theta)$ households. It is established that human capital along defines the learning growth (v_t) of school going children. Conditional on membership to Z_t for given $s_j(\theta)$ human capital is homogenous within Z_t cell so that $E[v(\Lambda_t)\,|\,\mathrm{Z}_t, e_t = e_{t+1} = 1] = v_{\mathrm{Z}(t)}$ applies.

Assumption 11 implies that sub- populations of $s_j(\theta)$ households with ($e_{lt} = e_{lt+1} = 1$) choices define sub-group constant average learning rates when these sub-groups share equal levels of Λ_t. This also means that taking expectations for given sub-populations over Z_t helps to isolate the effect of v_t even if we do not directly observe the actual levels of Λ_t. However, the variables Z_t do not determine the 'time-variability' of the stochastic shock π_t so that $E[\pi_t \mid Z_t,..]$ remains as a conditional expectation that can only be revealed by estimates based on the sample analogue to the expectation. Therefore, Assumption 11 implies that $E[v(\Lambda_t)\pi_t \mid Z_t, e_t = e_{t+1} = 1] = v_{Z(t)}E[\pi_t \mid Z_t, e_t = e_{t+1} = 1]$; observing households with consecutive $e_{lt} = e_{lt+1} = 1$ but identical Z_t implies that their children have experienced identical learning growth ($v_{Z(t)}$) even if there exists variability on the realized π_t. Together Assumptions 10 and 11 modify the moment conditions in expression (11) as follows:

$$Let \quad \Xi_{tE} = \left(Z_{tE}, \mathrm{B}_{tE}, e_t = e_{t+1} = 1\right) \quad and \quad \Xi_{t\Gamma} = \left(Z_{t\Gamma}, \mathrm{B}_{t\Gamma}, e_t = e_{t+1} = 0\right)$$

$$\begin{aligned}
&From \;\; (8) \;\; and \;\; (11): \quad Regime \;\; 1 \quad (education) \\
&\mathrm{E}\left[\Delta\left(\ln E_{t,t+1}\right)\middle|\Xi_{tE}\right] = \\
&\mathrm{E}\left[\Delta\left(\ln W_{t,t+1}\right)\middle|\Xi_{tE}\right] + \mathrm{E}\left[\Delta\left(\ln \pi_{t,t+1}\right)\middle|\Xi_{tE}\right] + v_{Z(t)}\mathrm{E}\left[\left(\pi_t\right)\middle|\Xi_{tE}\right] + \mathrm{E}\left[\kappa(\Lambda_t)\middle|\Xi_{tE}\right] \\
&From \;\; (10) \;\; and \;\; (11): \quad Regime \;\; 2 \quad (no \;\; education) \\
&\mathrm{E}\left[\Delta\left(\ln \Gamma_{t,t+1}\right)\middle|\Xi_{t\Gamma}\right] = \\
&\mathrm{E}\left[\Delta\left(\ln W_{t,t+1}\right)\middle|\Xi_{t\Gamma}\right] + \mathrm{E}\left[\Delta \ln\left(\pi + n^{1/2}\cdot\eta_{\bar{g}}\right)_{t,t+1}\middle|\Xi_{t\Gamma}\right] + \mathrm{E}\left[\kappa(\Lambda_t)\middle|\Xi_{t\Gamma}\right]
\end{aligned} \tag{12}$$

$$where \quad \mathrm{E}\left[\kappa(\Lambda_t)\middle|\Xi_{tE}\right] = \mathrm{E}\left[\kappa(\Lambda_t)\middle|\Xi_{t\Gamma}\right] \quad for \quad \Xi_{tE} \simeq \Xi_{t\Gamma} \quad \Rightarrow \left(\mathrm{B}_{tE} = \mathrm{B}_{t\Gamma}\right) \;\&\; \left(Z_{tE} \simeq Z_{t\Gamma}\right)$$

The two population moment conditions in (12) are the two conditions driving the empirical identification of the key human capital growth parameters $v(.)$ and $\kappa(.)$: except for these two parameters all terms in (12) explain conditional variability that once observed can be used to estimate the two parameters of interest for different (conditional) $s_j(\theta)$ sub-population. However, expression (12) is not operational at the empirical level because the innovations (π_t, π_{t+1}) are built within the model but not directly observed in the data, i.e., the expectations ($E[\pi_t \mid ..]$), ($\Delta \ln \pi_{t,t+1}$) and ($\Delta \ln\left(\pi + n^{1/2}\cdot\eta_{\bar{g}}\right)_{t,t+1}$) are not identified by the data. However, the consistency of our model is built on the assumption that household's schooling choices result directly from the household's evaluation of the innovation π_t (i.e., the change to the state of the world) against the household's schooling reservation policy π_t^R. Section 3.6 justifies that such reservation policy is a function of the household's preferences and characteristics that have been shaped conditional on the present and past investment history in the household. Lemma 1 implies such definition of π_t^R and shows that the structural model is built on the assumption that observing $e_{it} = 1$ at t implies a stochastic shock such that $\pi_{it} \geq \pi_{it}^R$, whereas $e_{it} = 0$ at t implies an evaluation of the world at t such that $\pi_{it} < \pi_{it}^R$. If we can retrieve π_t^R for each ith household this would provide a bound on the unknown magnitude π_{it} for given educational regime. At the same time, Assumption 1 implies a natural bound on the actual distribution from where the shocks are draws such that $\pi \in [\underline{\pi}, \overline{\pi}]$ for any ith household and at any $t \in [0, T]$ and, therefore, the following applies:

$$\begin{array}{l} For \quad \pi \in [\underline{\pi}, \overline{\pi}] \quad and \quad given \quad \pi_{it}^R = \pi_{it}^R\left(X', \psi_t^T\right),\; e_{it} \in \{0,1\} \; bounds \; \pi_{it}: \\ \Rightarrow \\ Regime\ 1 \quad (education): \qquad e_{it} = 1 \quad \Rightarrow \quad \pi_{it}^R \le \pi_{it} \le \overline{\pi} \\ Regime\ 2 \quad (no\ education): \qquad e_{it} = 0 \quad \Rightarrow \quad \underline{\pi} \le \pi_{it} \le \pi_{it}^R \end{array} \tag{13}$$

Condition (13) combines Assumption 1 and Lemma1 and bounds the unknown magnitude of the stochastic shock π_{it} providing a links observed variability to unobserved stochastic shocks for given educational regime. The bounding set $[\underline{\pi}, \overline{\pi}]$ is a normalizing assumption necessary for the identification of the value function in (5) as shown in Appendix 1 so that the bounds in (13) are consistent with our modelling strategy. The set of bounds in (13) further implies that the three expectations in (12) – that is, $E[\pi_t \mid ..]$, $\Delta \ln \pi_{t,t+1}$ and ($\Delta \ln\left(\pi + n^{1/2} \cdot \eta_{\overline{g}}\right)_{t,t+1}$) – can be identified up to a bounding interval if and only if the we can use observed household's behaviour to elicit information on π_{it}^R and the two limiting values $[\underline{\pi}, \overline{\pi}]$ that are assumed to be common to all members in the population (see Assumption 1) and therefore independent from idiosyncratic household characteristics or even time period. By behaviour we mean observed choice behaviour that at a population level implies a probability for given household types conditional on preferences and characteristics. That is, in population terms we may define (13) in terms of probabilities; first, all i -households such that $e_{it} = 1 \Rightarrow \pi_{it}^R \le \pi_{it} \le \overline{\pi}$ are associated with the population conditions $P(e_{it} = 1) = P(\pi_{it}^R \le \pi_{it} \le \overline{\pi})$, whereas the alternative choice $e_{it} = 0 \Rightarrow \underline{\pi} \le \pi_{it} \le \pi_{it}^R$ implies the alternative population condition $P(e_{it} = 0) = P(\underline{\pi} \le \pi_{it} \le \pi_{it}^R)$. These two probabilities may become empirically operational if we can relate the two population probabilities ($P(e_{it} = 0)$ and $P(e_{it} = 1)$) to the household's specific preferences and characteristics. Throughout the paper we have associated the reservation policy $\pi_{it}^R = \pi_{it}^R\left(X', \psi_t^T\right)$ to such household's specifics. Let K_t be a set of variables that define such preferences and characteristics.[35] With this, the link between observed variability and the

[35] These will include all those variables at household level that may define the fixed cost of schooling for the household since these fixed costs are the ones that ultimately define the reservation policy for the schooling choice. The set may include household characteristics such as wealth variables (e.g., material of household), household member's characteristics and number of household members, regional characteristics, available public and private services to the household, level of urbanization and pricing information if this is heterogeneous among distinct household types.

schooling reservation policy is given by $\pi_{it}^R = \pi_{it}^R(X'_{it}, \psi_t^T) = \pi_{it}^R(\mathrm{K}_{it})$. According to Assumption 1 and Lemma 1 the consistency of the model requires a positive support of the value function so that $0 \le \underline{\pi} \le \pi_{it}^R \le \overline{\pi} < \infty$ is required. Therefore, whatever way we choose to characterise $\pi_{it}^R = \pi_{it}^R(\mathrm{K}_{it})$ this has to be so that π_{it}^R is always positive. An easy way to achieve such requirement is to let π_{it}^R be an exponential function of the set K_t with Assumption 3 guaranteeing that $\pi_{it}^R < \infty$ even if we assume an exponential representation for π_t^R.[36] In what follows Assumption 13 provides the representation for the reservation policy:

Assumption 13 (characterization of education reservation policy):

$$\begin{aligned} &\textit{Let} \quad \mathrm{K}_t \ \textit{explain Preferneces and Characteristics}: \\ &\Rightarrow \qquad \pi_{it}^R = \pi_{it}^R(\mathrm{K}_{it}) = \exp(\mathrm{K}'_{it}\gamma) \\ &\Rightarrow \qquad \ln \pi_{it}^R = \mathrm{K}'_{it}\gamma \end{aligned} \tag{14}$$

The variables in K_{it} can vary by *ith* household but also by educational choices so that alternatively we may define $\mathrm{K}_{it:E}$ and $\mathrm{K}_{it:\Gamma}$. However, this is more of an empirical problem where the available data sets will help us to define the variables in K_{it}.[37] The vector of parameters (γ) in (14) weights each of the characteristic in K_t according how important these variables are with regards to the educational choice in the household. Therefore, using (14) and conditional on K_t, the two behavioural conditions $P(e_{it} = 1 | \mathrm{K}_{it})$ and $P(e_{it} = 0 | \mathrm{K}_{it})$ may also be expressed conditional on observable K_t as $P(e_{it} = 1 | \mathrm{K}_{it}) = P(\exp(\mathrm{K}'_{it}\gamma) \le \pi_{it} \le \overline{\pi})$ and $P(e_{it} = 0 | \mathrm{K}_{it}) = P(\underline{\pi} \le \pi_{it} \le \exp(\mathrm{K}'_{it}\gamma))$, respectively. Moreover, logarithmic transformations imply monotonic transformations so that the relation $0 \le \underline{\pi} \le \pi_{it}^R \le \overline{\pi} < \infty$ also holds in logarithms ($0 \le \ln \underline{\pi} \le \ln \pi_{it}^R \le \ln \overline{\pi} < \infty$) and this carries through to the population conditions, that is, $P(e_{it} = 1 | \mathrm{K}_t) = P(\ln \pi_{it}^R \le \ln \pi_{it} \le \ln \overline{\pi}) = P(\mathrm{K}'_{it}\gamma \le \ln \pi_{it} \le \ln \overline{\pi} | \mathrm{K}_t)$, while $P(e_{it} = 0 | \mathrm{K}_{it}) = P(\underline{\pi} \le \pi_{it} \le \pi_{it}^R) = P(\ln \underline{\pi} \le \ln \pi_{it} \le \mathrm{K}'_{it}\gamma | \mathrm{K}_{it})$. The logarithmic transformation of the support for the value function provides a tractable format so that the assumptions behind the model lead to an estimation procedure based on index models. Starting from (4), household's behaviour is justified as driven by utility maximization. Although we cannot directly observe the latent process that quantifies the utility obtained from alternative choice, the actual observed choice is the one that implies the highest level

[36] That is, Assumption 3 requires the set K_t to be convex, bounded and continuous. These requirements may be achieved if the set K_t is finite in variables and the variables imply realizations well defined in a finite set.

of utility conditional on observed behaviour. But this is exactly the same behaviour reasoning that justifies the use of conditional binary outcomes to estimate the effect of covariates on observed behaviour.[38] If so, we use the logarithmic transformation to provide a (conditional) binary interpretation as follows:

Let $e_{it}^{*} = f(\mathrm{K}_{it};\gamma) + \varepsilon_{it}$ *be the unobserved latent process that explains the following observational rule applies*;

$$
\begin{aligned}
&(i)\quad \textit{Observing } e_{it} = 1 \quad \Rightarrow \quad e_{it}^{*} \geq 0 \quad \Rightarrow \quad \textit{positive utility from schooling} \\
&\qquad\qquad\qquad\qquad\qquad\quad \Rightarrow \quad f(\mathrm{K}_{it};\gamma) + \varepsilon_{it} \geq 0 \\
&(ii)\quad \textit{Observing } e_{it} = 0 \quad \Rightarrow \quad e_{it}^{*} < 0 \quad \Rightarrow \quad \textit{negative utility from schooling} \\
&\qquad\qquad\qquad\qquad\qquad\quad \Rightarrow \quad f(\mathrm{K}_{it};\gamma) + \varepsilon_{it} < 0
\end{aligned}
$$

Let $f(\mathrm{K}_{it};\gamma)$ *be given as* $f(\mathrm{K}_{it};\gamma) = \mathrm{K}_{it}'\gamma$

From (13),

$$
\begin{aligned}
&\textit{Education}: \quad e_{it} = 1 \quad \Rightarrow \quad \pi_{it} \geq \pi_{it}^{R} \quad \Rightarrow \quad \ln \pi_{it} \geq \ln \pi_{it}^{R} \quad \Rightarrow \quad \ln \pi_{it} \geq \mathrm{K}_{it}'\gamma \\
&\qquad\qquad\qquad\qquad\quad \Rightarrow \quad \pi_{it} \geq \pi_{it}^{R} \quad \Rightarrow \quad \ln \pi_{it} \geq \ln \pi_{it}^{R} \quad \Rightarrow \quad \ln \pi_{it} \geq \mathrm{K}_{it}'\gamma \\
&(15.i) \qquad P(e_{it}^{*} \geq 0 \mid \mathrm{K}_{it}) = P(e_{it} = 1 \mid \mathrm{K}_{it}) = P(\ln \overline{\pi} \geq \ln \pi_{it} \geq \mathrm{K}_{it}'\gamma \mid \mathrm{K}_{it}) \\
&\\
&\textit{No-education}: \quad e_{it} = 0 \quad \Rightarrow \quad \pi_{it} < \pi_{it}^{R} \quad \Rightarrow \quad \ln \pi_{it} < \ln \pi_{it}^{R} \quad \Rightarrow \quad \ln \pi_{it} < \mathrm{K}_{it}'\gamma \\
&\qquad\qquad\qquad\qquad\quad \Rightarrow \quad \pi_{it} < \pi_{it}^{R} \quad \Rightarrow \quad \ln \pi_{it} < \ln \pi_{it}^{R} \quad \Rightarrow \quad \ln \pi_{it} < \mathrm{K}_{it}'\gamma \\
&(15.ii) \qquad P(e_{it}^{*} < 0 \mid \mathrm{K}_{it}) = P(e_{it} = 0 \mid \mathrm{K}_{it}) = P(\ln \underline{\pi} \leq \ln \pi_{it} < \mathrm{K}_{it}'\gamma \mid \mathrm{K}_{it})
\end{aligned}
\tag{15}
$$

For all, $\ln \pi_{it}$ *is an stochastic realization from some know distribution.*

In expression (15) the two conditional moments $(15.i)$ and $(15.ii)$ provide functional forms for the estimation of the parameter vector γ that may differ between educational sub-groups (see Section 4.3). For example, let $\ln \pi \sim N[0,\sigma]$ alongside the assumption $[\ln \underline{\pi}, \ln \overline{\pi}] = [-\infty, +\infty]$ so that $[\underline{\pi}, \overline{\pi}] = (0, +\infty)$. Assuming that we can identify the sub-population implied by e_{it} and the data is informative about the variables K_t we can estimate the index models $(15.i)$ and $(15.ii)$ using Probit specifications to obtain the parameter vectors $\hat{\gamma}_{t,E}$ and $\hat{\gamma}_{t,\Gamma}$. Having estimated these parameters, we can consistently project the otherwise unobserved reservation value $\hat{\pi}_{it}^{R} = \exp(\mathrm{K}_{it}'\hat{\gamma}_t)$ for given education sub-samples. However,

---

[37] That is, we do not need specific identification strategies for π_t^R according to educational regime, since the household characteristics K_t already shape the reservation policy and distinguishes households by regime choice once π_t^R is determined.

[38] For example, the choice between working or not in well defined labour markets and conditional on all those variables that determine the fixed cost of work; in developed economies it is common to summarize these fixed cost with a all variables that may explain the individual's reservation wage.

using expression (13) we see that the assumption $[\underline{\pi},\overline{\pi}]=(0,+\infty)$ is not realistic (or acceptable by the requirements of the model) or even sufficient to narrow the interval in (13) so to pin down the magnitude of the unobserved innovation π_{it}. Section 4.3 below suggests an estimation procedure to retrieve the set $[\underline{\pi},\overline{\pi}]$ from the data. However, for given values $[\underline{\pi},\overline{\pi}]$ the parameters $v(.)$ and $\kappa(.)$ require characterizing the (assumed) known distribution from which the innovations draw commonly for all in the population. Assumption 14 characterises such distribution thus completing Assumption 1:

Assumption 14 (common distribution on innovations): The productivity innovation π follows a truncated lognormal distribution $\ln N\left(\mu,\sigma_{\pi}\right)$ in the support $[\underline{\pi},\overline{\pi}]$ where $\underline{\pi}<\overline{\pi}$ for $[\underline{\pi},\overline{\pi}]\in\left(0,\infty\right)$. This implies that $\ln\pi$ follows the truncated normal distribution $N\left(\mu_{\ln\pi},\sigma_{\ln\pi}\right)$ in the support $[\ln\underline{\pi},\ln\overline{\pi}]$ where we assume the normalization $\mu_{\ln\pi}$ =0.

Assumption 14 provides an interpretation of a distribution common to all household in the population from where the productivity shocks are drawn. The assumption is justified: a log normal distribution for $[\underline{\pi},\overline{\pi}]$ would imply that on average households receive low to average stochastic shocks that leaves choices relatively unaltered while innovations from the extremes of the distribution (very low bad shocks and extremely positive good shocks) are possible but less likely for the population as a whole. The practicality of allowing that π follows a log-Normal distribution (so that its log transformation follows the Normal distribution) becomes obvious when presenting the estimation procedure in Section 4.3. Whereas Assumptions 1 to 9 provide the backing for Lemmas 1 to 3 to provide the consistency of the model, Assumptions 10 to 14 provide the full set of conditions that allow for the identification of two human capital growth parameters in the presence of the appropriate data set. Section 4.3 below provides the estimation procedure required as elicit the set parameters $v(.)$ and $\kappa(.)$ from the data.

4.3 Estimation method[39]

Assume that our data set is sufficiently informative so that we may observe individuals in specific educational regimes (i.e., $e_{it}=1$ or $e_{it}=0$) as well as observing all variables that may constitute the set K_{it} (i.e., household's i preferences and characteristics at time t). Our model suggest that the education choice maximises utility for the household subject to the normally distributed stochastic variability (i.e., $\ln\pi_{it}\sim N(0,\sigma_{\ln\pi})$ that underlines an unobserved latent utility valuation of the state of the world. Imposing normality implies that we can consistently estimate γ_t using a probit specification where the dependent

[39] An alternative estimation procedure with higher parametric content will be provided in Appendix 3. However, the estimation procedure proposed in this section aims at an estimation method that is as flexible as possible and to this aim it is similar to that put forward in Lechner and Vazquez-Alvarez (2007).

variable is the binary outcomes $e_{it}=1$ (against $e_{it}=0$) conditional on K_{it}. However, unlike the classic probit set up, we need to transform the density (and therefore the probit specification) so that it takes into account the truncation dictated by the upper and lower bounds of the support for the value function, i.e., $[\underline{\pi},\overline{\pi}]$. Allowing for such truncation we define a time dependent likelihood function consistent with the behavioural assumptions in the model:

$$\begin{array}{ll} Regime\ 1\quad (education): & e_{it}^{*}\geq 0 \ \Rightarrow\ e_{it}=1 \ \Rightarrow\ P(e_t=1\,|\,\mathrm{K}_t;\gamma) \\ Regime\ 2\quad (no\ \ education): & e_{it}^{*}<0 \ \Rightarrow\ e_{it}=0 \ \Rightarrow\ P(e_t=0\,|\,\mathrm{K}_t;\gamma) \end{array}$$

$$\begin{array}{l} where\quad P(e_t=1\,|\,\mathrm{K}_t;\gamma)=P(\ln\pi_t^{R}\leq\ln\pi_t\leq\ln\overline{\pi}\,|\,\mathrm{K}_t;\gamma)=P(\mathrm{K}_t'\gamma\leq\ln\pi_t\leq\ln\overline{\pi}\,|\,\mathrm{K}_t;\gamma) \\ and\quad\ \ P(e_t=0\,|\,\mathrm{K}_t;\gamma)=1-P(e_t=1\,|\,\mathrm{K}_t;\gamma) \\ \Rightarrow \\ P(e_t=1\,|\,\mathrm{K}_t;\gamma)=\left(\dfrac{\Phi(\ln\overline{\pi})-\Phi(\mathrm{K}_t'\gamma)}{\Phi(\ln\overline{\pi})-\Phi(\ln\underline{\pi})}\right)\quad\&\quad P(e_t=0\,|\,\mathrm{K}_t;\gamma)=\left(\dfrac{\Phi(\mathrm{K}_t'\gamma)-\Phi(\ln\underline{\pi})}{\Phi(\ln\overline{\pi})-\Phi(\ln\underline{\pi})}\right) \\ where\ \ \ln\pi_t\sim N(0,\sigma_{\ln\pi})\quad\Rightarrow\Phi: Normal\ cdf \\ \Rightarrow \\ L(\gamma\,|\,\underline{\pi},\overline{\pi},\mathrm{K}_t)=\prod_{i\in N}\left(\dfrac{\Phi(\ln\overline{\pi})-\Phi(\mathrm{K}_{it}'\gamma)}{\Phi(\ln\overline{\pi})-\Phi(\ln\underline{\pi})}\right)^{I(e_{it}=1)}\left(\dfrac{\Phi(\mathrm{K}_{it}'\gamma)-\Phi(\ln\underline{\pi})}{\Phi(\ln\overline{\pi})-\Phi(\ln\underline{\pi})}\right)^{I(e_{it}=0)} \end{array} \tag{16}$$

The maximum likelihood estimate $\hat{\gamma}_{MLE,t}$ is conditional both on the variable's set K_t at t and the pair of values $[\underline{\pi},\overline{\pi}]$ where both K_t and $[\underline{\pi},\overline{\pi}]$ are defined for $i=1,...,N$ households representing an independent random sample from some given population. We notice that the likelihood function may also be defined by $L(\gamma,\underline{\pi},\overline{\pi}\,|\,\mathrm{K}_t)$ so that the two values $[\underline{\pi},\overline{\pi}]$ are included in the parameter space. Such an alternative to the likelihood function in (16) is correct and may be implemented for some set of starting up values $[\underline{\pi},\overline{\pi}]$. However, as was the case in Lechner and Vazquez-Alvarez (2007), the cross derivatives $(\partial/\partial\gamma\partial\underline{\pi})$, $(\partial/\partial\gamma\partial\overline{\pi})$, $(\partial/\partial\overline{\pi}\partial\underline{\pi})$ and $(\partial/\partial\underline{\pi}\partial\overline{\pi})$ enter as off diagonals in the information matrix. Whereas in expectations $E(\partial/\partial\gamma\partial\underline{\pi})=E(\partial/\partial\gamma\partial\overline{\pi})=0$, the functions $(\partial/\partial\overline{\pi}\partial\underline{\pi})$ and $(\partial/\partial\underline{\pi}\partial\overline{\pi})$, and their higher order derivatives enter the Hessian in multiplicatively with the probability density function $\phi(\ln\pi)$ and $\phi(\ln\pi)$ can be multiplicatively close to zero for small changes in the values $[\underline{\pi},\overline{\pi}]$ as would happen when the iteration process approaches the maximum value of the likelihood function. The result is the singularity of the information matrix at which point the likelihood function is not identified. However, we know that the values $[\underline{\pi},\overline{\pi}]$ exists when the problem is well defined and that such values are fixed and unique for all households in the population since all households share the same distribution of innovations. For example, we know that for some fixed values $[\underline{\pi},\overline{\pi}]$: $\underline{\pi}\simeq 0,\overline{\pi}\simeq\infty$ the values

$[\ln \underline{\pi}, \ln \overline{\pi}]$: $\ln \underline{\pi} = -\infty, \ln \overline{\pi} = \infty$ and the likelihood function in (16) falls down to the classic likelihood function without truncation: in fact, the pair $[\underline{\pi}, \overline{\pi}]$ only acts as systematically shifting the likelihood function but leaves the property of the maximum likelihood for the vector γ unchanged. In our case, if we knew the set $[\underline{\pi}, \overline{\pi}]$ we could just insert these into the likelihood function in (16) and estimate the parameter vector $\hat{\gamma}_{MLE,t} \,|\, [\underline{\pi}, \overline{\pi}]$ by probit as implied by $\ln \pi_{it} \sim N(0, \sigma_{\ln \pi})$. But for unknown $[\underline{\pi}, \overline{\pi}]$ the correct procedure is to fix the set $[\underline{\pi}, \overline{\pi}]$ and estimate the likelihood function for different combinations of these and select the pair $[\underline{\pi}^*, \overline{\pi}^*]$ that implies the highest value of the likelihood function so that our final likelihood estimates are given by $\left(\hat{\gamma}_{MLE,t} \,|\, [\underline{\pi}^*, \overline{\pi}^*], \mathrm{K}_t\right) = \max_{[\underline{\pi}, \overline{\pi}] \in [0,b]} \left(L(\gamma \,|\, \underline{\pi}, \overline{\pi}, \mathrm{K}_t)\right)$.[40] Having obtained the vector $(\hat{\gamma}_{MLE,t}, \underline{\pi}^*, \overline{\pi}^*)$ from the estimation procedure, the weights $\hat{\gamma}_{MLE,t}$ imply retrieving the reservation policy $\hat{\pi}^R_{it} = \exp(\mathrm{K}'_{it} \hat{\gamma}_{MLE,t})$ for each $i \in N$ in the sample and, together with the values $[\underline{\pi}^*, \overline{\pi}^*]$, we may build a bounding interval on the unknown magnitudes of the stochastic innovation as expressed in (13). The intervals may be used in two distinct forms. On the one hand, assuming that all values in the interval are equally likely we may take the midpoint of the interval (let this midpoint be $\widehat{\pi}_{it}$) to approximate π_{it}: if so, the sample analogues to $\mathrm{E}[\pi_t \,|\, \Xi_{t,E}]$, $\mathrm{E}[\pi_t \,|\, \Xi_{t,E}]$, $\mathrm{E}\left[\Delta(\ln \pi_{t,t+1}) \middle| \Xi_{tE}\right]$ and $\mathrm{E}\left[\Delta \ln(\pi + n^{1/2} \cdot \eta_{\bar{g}})_{t,t+1} \middle| \Xi_{t\Gamma}\right]$ can be approximated by $(1/N)\sum_{i=1}^{N} f(\widehat{\pi}_{it} \,|\, \Xi_{t,\omega}), \omega = E, \Gamma$ where $f(\cdot)$ implies any of the four expectations. With this, values $v(.)$ and $\kappa(.)$ can be estimated with precision as point estimates from the sample analogue to each of the corresponding expressions in (12). An alternative is to approximate each of the four expectations with the sample analogue based on the actual values of the bounding interval (e.g., $\mathrm{E}[\pi_t \,|\, \Xi_{t,E}]$ may be approximated by the estimable interval $\left\{(1/N)\sum_{i=1}^{N} (\hat{\pi}^R_{it} \,|\, \Xi_{it,E})\right\} \le \mathrm{E}[\pi_{it} \,|\, \Xi_{it,E}] \le \left\{(1/N)\sum_{i=1}^{N} (\overline{\pi}^* \,|\, \Xi_{it,E})\right\} = \overline{\pi}^*$) but in this case the parameters $v(.)$ and $\kappa(.)$ would only be identified up to an interval implied by the expectations and expressions in (12). Assuming that we apply the point estimate procedure implied by the midpoint, the following shows the final sample analogue to each of the expressions in (12):

[40] In practice we start by setting the values of $\underline{\pi} \in [0,u]$ and $\overline{\pi} \in [0,u]$ for some value of u that implies a maximum on the likelihood function (e.g., letting $u \simeq 30$, $\Phi(\ln \overline{\pi}) \simeq 1$). These bounds imply a grid of pair wise combinations and we estimate the value of the likelihood function for all these combinations. We select the vector $(\hat{\gamma}_{MLE,t} \,|\, [\underline{\pi}^*, \overline{\pi}^*])$ so that the pair $[\underline{\pi}^*, \overline{\pi}^*]$ imply the highest value of the likelihood function among all estimated likelihood functions within the defined grid. Notice that identification relies on $u < 30$ otherwise identification is not achieved and alternative methods would be required – see Appendix 3 – where the estimation procedure would not rely on the properties implied by maximum likelihood estimation. A further refinement to estimate the pair $[\underline{\pi}^*, \overline{\pi}^*]$ would involve different waves so that $([\underline{\pi}^{**}, \overline{\pi}^{**}]) = [\min_t \{\underline{\pi}^*_1, \underline{\pi}^*_2, ..., \underline{\pi}^*_t\}, \max_t \{\overline{\pi}^*_1, \overline{\pi}^*_2, ..., \overline{\pi}^*_t\}]$. However, this last estimation procedure would imply the assumption that between time periods the distribution of innovations ‘common to all households’ remains constant over time, an assumption that may not be realistic.

$$Let \quad \Xi_{tE} = \left(Z_{tE}, \mathrm{B}_{tE}, e_t = e_{t+1} = 1\right) \quad and \quad \Xi_{t\Gamma} = \left(Z_{t\Gamma}, \mathrm{B}_{t\Gamma}, e_t = e_{t+1} = 0\right)$$

$$For \;\; a \;\; sub-population \;\; s_j(\theta), \;\; j = \{1,2,3,4\}, \;\; \theta \in \left\{\theta_1, \theta_2, ..., \overline{\theta}\right\}$$

$$From \;\; (12): \quad Regime \;\; 1 \quad (education)$$

$$\hat{v}\left(s_j(\theta) \,|\, \Xi_{tE}\right) = \frac{\left(\dfrac{1}{N \in \Xi_{tE}}\right)\left(\sum_{i=1}^{N \in \Xi_{tE}} \Delta \ln E_{it,t+1}^{s_j(\theta)} - \sum_{i=1}^{N \in \Xi_{tE}} \Delta \ln W_{s_j(\theta)t,t+1} - \sum_{i=1}^{N \in \Xi_{tE}} \Delta \ln \widehat{\pi}_{t,t+1} - \hat{\kappa}\left(s_j(\theta) \,|\, \Xi_{tE}\right)\right)}{\left(\dfrac{1}{N \in \Xi_{tE}}\right)\left(\sum_{i=1}^{N \in \Xi_{tE}} \widehat{\pi}_{it}\right)}$$

$$From \;\; (12): \quad Regime \;\; 2 \quad (no \;\; education)$$

$$\hat{\kappa}\left(s_j(\theta) \,|\, \Xi_{t\Gamma}\right) = \left(\frac{1}{N \in \Xi_{t\Gamma}}\right)\left(\sum_{i=1}^{N \in \Xi_{t\Gamma}} \Delta \ln \Gamma_{it,t+1}^{s_j(\theta)} - \sum_{i=1}^{N \in \Xi_{t\Gamma}} \Delta \ln W_{s_j(\theta)t,t+1} - \sum_{i=1}^{N \in \Xi_{t\Gamma}} \Delta \ln\left(\widehat{\pi} + n^{1/2} \cdot \eta_{\overline{g}}\right)_{t,t+1}\right)$$

$$where \quad \hat{\kappa}\left(s_j(\theta) \,|\, \Xi_{t\Gamma}\right) = \hat{\kappa}\left(s_j(\theta) \,|\, \Xi_{tE}\right), \;\; for \quad \Xi_{tE} \simeq \Xi_{t\Gamma} \quad \Rightarrow \left(\mathrm{B}_{tE} = \mathrm{B}_{t\Gamma}\right) \;\&\; \left(Z_{tE} \simeq Z_{t\Gamma}\right) \qquad (17)$$

Thus, estimation of the two human capital growth parameters implies population sub-cell estimation where these sub-cells are defined by variables in Ξ_{tE} and $\Xi_{t\Gamma}$ that further condition our estimates to observing the same household over two consecutive time periods, that is, identification relies on observing households such that $e_{it} = e_{it+1} = 1$ and $e_{it} = e_{it+1} = 0$ where the distance between t and $t+1$ will be dictated by the nature of the panel structure. The estimates for estimation therefore rely on (a) sorting of the data by sample sub-groups $s_j(\theta)$ where each member of the subgroup needs to be observed over (at least) to consecutive time periods, identification of the set $(\pi_t^R, \underline{\pi}^*, \overline{\pi}^*)$ using a truncated likelihood function whose likelihood estimates relies on a probit specification and finally, estimation of the value parameters in (17) implies simply calculating these with each of the constructed cells. Clearly, our flexible semi-parametric procedure relies on the availability of variables for Z_t and B_t that provide sufficient information for the purpose of identification.

At this point we deal with the tractability of $\eta_{\overline{g}}$ that emulates the likeness of children relative to adults as productive units in the work place. A simple specification is that of letting $\eta_{i\overline{g}} = (\overline{g}(i,t)/18)$, where the term $\overline{g}(i,t)$ stands for the average age of children in household i at time t: this simple specification is sufficient to capture that as children get closer to adults ($\overline{g}(i,t) \rightarrow 18$) their capacity to emulate these in the workplace becomes 1. A final point to make is that of making inference from the sample to the population: the two parameters in (17) require the intermediate step of estimating the weights $\hat{\gamma}_{MLE,t}$ and this intermediate step may affect the variance-covariance matrix associated with the two parameters. To avoid possible nuisance effects from the first step on the estimates of the sampling variance the alternative is to use some bootstrap procedure to emulate the empirical distribution of the sampling error, for example, the use of a naïve bootstrap procedure thus re-sampling randomly with replacement from the data.

5. The Kagera Health and Development Survey (1991-2004)

The Kagera Health and Development Survey (KHDS, 1991-1994 and 2004) is a research initiative on behalf of the Development Research group in the World Bank. The survey provides empirically based evidence on the interrelation between health and economic outcomes – growth and poverty - in the long run. The survey is representative of Kagera, a region in the North West of Tanzania with approximately 2 million inhabitants (circa 2000 census) close to Lake Victoria, bordering with Uganda and Rwanda and near South-West Kenya and North-East Burundi. The KHDS survey started in September 1991 and collected socio-economic information over a period of 29 months – i.e., until January 1994, inclusive – from households representing 51 regional districts. During the survey period each responding household was interviewed at least once and at most up to four times with a gap of approximately 6-7 months between each of the interviews. Following up from the last survey period (January 1994) the study was further complemented with a final interview conducted during 2004, thus providing complementary information to help understand the socio-economic dynamic information from the same households over (at most) 13 years. The initial number of interviewed households (first round, September 1991 to May 1992) was 843. The rate of attrition from the first to the second round (April 1992 to November 1992) was 4.4; a further 46 new households were added to the sample during the second round to replace the initial loss due to attrition. Likewise, a 2.2% attrition between the second and third round (November 1992 to May 1993) was further accounted for with the addition of 30 new replacement households, whereas in the forth round (June 1993 to January 1994) only remaining households from the third round were interviewed. In total, 919 households were survey at least once between September 1991 and January 1994 and altogether these households represented a total of 6,681 individuals surveyed at least ones over the period. The age of these 6,681 individuals ranged from 0 to 107 with answers for youngsters and infants in the household provided by the household respondent who was also responsible to answer questions at household level (e.g., durables, household's business, etc.). Expression (17) shows the two key parameters that we aim at estimating and how the dynamics of these rely on changes in productivity gains conditional on the observed educational choices in the household. It is clear from (17) that estimating such dynamics requires two crucial sample selection criteria; first, households have to be observed consecutively at least over two time periods (t and $t+1$) and secondly we are only interested on households with children in the age range 6 to 15, i.e., households that have to make schooling decisions for the primary school-age children in the household.[41] Table 1 shows the dynamic changes that sample units (number of households, individuals and children) experienced between the first survey round in September 1991 and the last survey round in January 1994.[42]

Table 1: Sample dynamics (September 1991 – January 2004)

	First Survey (P1) [Sep 1991 – May 1992]	Second Survey (P2) [April 1992 – Nov 1992]	Third Survey (P3) [Nov 1992 – May 1993]	Forth Survey (P4) [June 1993 – Jan 1994]
Number of Households	843	880	878	852
Without children in [6,15] age range (T1)	211	210	182	156
With children in [6,15] age range (T2)	632	670	696	696
Cumulative Number of households if T2(1)				
Conditional on positive number of kids at P1	632	628	612	602
With positive number of kids at all P's P1-P4	632	620	596	579
Number of Individuals if T2 (2)	4,259	4,804	5,400	5,762
Cumulative Number of individuals if T2 (3)	4,249	4,502	4,743	4,936
Number of children age range [6,15] in cumulative households if T2 (3)	1,583	1,652	1,697	1,691
Number of boys age range [6, 15] in cumulative households if T2 (3)	805	825	834	834
Number of girls age range [6, 15] in cumulative households if T2(3)	778	827	863	857
Weighted Average Household size if T2 (3)	6.7 (3.0)	7.3 (3.2)	8.0 (3.5)	8.5 (3.7)

Note: Source, own estimates KHDS 1991-1994. (1) T1 (Type 1) households are those without primary school age children and T2 (Type 2) households are those with at least one child in the age range 6 to 15 at the time of the survey, i.e., school age up to the last year in primary education. (2) Refers to households that are observed at each point in time (each Pt, t=1,2,3,4) with a positive number of children; these households are of size 632, 670, 696 and 696 for P1, P2, P3 and P4 respectively (see Row 1, line 3). (3) Refers to households that are observed over subsequent periods and at each period implies a positive number of children; these households are of sizes 632, 620, 596 and 579 for P1, P2, P3 and P4 respectively (see Row 2, line 2 in this same table). The average household size is based on weighted averages using the survey sample weights; bracketed numbers show standard deviations.

From Table 1 we see that only 632 of the 843 initially surveyed households contain children of primary school age during the first round. Of these, only 579 are consistently observed throughout the study period with the remaining 53 households dropping at some point before reaching the 4th round of the

[41] School in Tanzania starts at the age 6 (or 7) and the system is such that children are expected to complete primary school approximately at the age of 15 (or 16). Secondary education starts from the age of 16. Whereas we could consider educational decisions of the household beyond primary school (i.e., for children in the household with ages between 16 and 19), we consider that educational decisions at levels higher than primary level imply choices associated with states of the world that are not a concerned to our modeling assumptions (i.e., once the choice is between increasing levels of secondary education, shocks to household's productivity may be less relevant than other effects – e.g., government subsidies). Likewise, it is the case that once primary education is completed, we may think that with reference to the average in the population, the child has reached the flat portion of the human capital schedule. For these reasons we focus our estimation to the educational choices of children at primary education age.

[42] Although further information was collected in 2004, respondents in 2004 were the same individuals that were ever sampled between 1991 and 1994. Without any new households added in 1994 (neither in 2004) our sample in the first four rounds is identical to that in 2004. As such, the information in 2004 is static for that period and provides socio-economic outcomes for the sample 10 years ahead from the last round in 1994. This means that only information from 1991 to 1994 implies dynamic changes that concerns the educational decision of the 919 households ever surveyed. Therefore, at this stage of our analysis we rely solely on information collected during the 1st, 2nd, 3rd and 4th rounds. Whereas household weights are used throughout the empirical section to make inference from the sample to the population, the weights are not relevant for Table 1 because such table is only intended to inform the reader about the sample size and sample size movements over the period of study.

survey. Initially, the 632 households represent 4,259 individuals of which 1,583 are primary school age children (805 boys and 778 girls). Over time the size and composition of these households change as result of aging children, new births and mortality. The result is that in the final round the 579 consecutively observed households represent a total of 4,936 individuals of which 1,681 are primary school age children (834 boys and 857 girls). Table 1 puts emphasis on distinguishing among adults and children because according to our model it is the children who eventually have the potential to drive the human capital formation in the household.

The sample selection requirement that households need to be observed consecutively over time means that we could focus on the 579 consecutively observed households as shown in Table 1: this would imply defining the distance between t and $t+1$ to be approximately 18 months, and implicitly we would be assuming that educational choices are made every 18 months. However, it is not unreasonable to assume that in developing economies households make educational choices more often than once per year: children who start attending school in September may be required to stop attendance before completing the current grade if the household requires them for the purpose of work and (or) the household suffers adverse productivity shocks: such reality implies that the distance between t and $t+1$ is best measured at intervals shorter than 18 months if our aim is to capture the effects of productivity shocks on educational decisions in the household. Starting from the first round – September 1991 to May 1992 – the household's respondent is asked to declare the schooling status of all the school age children in the household at each and every round; the questions are such that we are able to capture the schooling status of the child without confounding effects that could occur if the survey falls during a public school holiday.[43] We take advantage of such detailed information on the dynamics of schooling choices to define the distance between t and $t+1$ in expression (17) as that implied between rounds, i.e., approximately six to seven months. Based on this, estimates of v_t and κ_t as explained in (17) reflect growth rates on human capital between three possible sets of consecutive time periods: $P1$ and $P2$, $P2$ and $P3$, and finally, $P3$ and $P4$. It is important to notice that the growth parameters (i.e., v_t and κ_t for schooling choices, or κ_t for non-schooling choices) are in fact the curvatures of a concave human capital function (i.e., $\Lambda_{t+1} = \Lambda_t \exp(v_t)\exp(\kappa_t)$ or $\Lambda_{t+1} = \Lambda_t \exp(\kappa_t)$, respectively). For this reason a household that is observed over distinct consecutive periods count as distinct consecutive observations. An example will be sufficient to make this point more clear: assume that household l is observed over the periods $P1$, $P2$ and $P3$ but drops thereafter so that it is no longer observed at $P4$. Such household provides consecutive dynamic

[43] The actual question being '*is [name of child] attending school now?*' followed by questions that aim at eliciting the characteristics of the school (e.g., private or public, hours attended, etc). When the answer to the question regarding current attendance is 'no', in the case of the first round (September 1991 – May 1992) the respondent is asked '*has [name of child] attended school during the past 12 moths*?', whereas in subsequent rounds the question is identical but with reference to '*6* (and not *12*) *months*', thus reflecting the elapse of time between rounds.

information on its possibly changing income, changing productivity and changing schooling choices between two distinct consecutive time periods, namely, $P1$ and $P2$ as well as $P2$ and $P3$. Any dynamic productivity change experienced by household l over the full observation spell – i.e. from $P1$ to $P3$ – implies cumulative growth rates between the two subsequent periods i.e., growth rate during $P1$ and $P2$ followed by growth rates during $P2$ and $P3$. Because our interest is on the growth rate, a household that is observed subsequently over distinct sets of time periods counts as different '*observed units*' that experience distinct changes in human capital. The only requirement for this estimation strategy to work is that all monetary values (i.e., income values) have to be converted into real terms. Thus, growth rates imply real growth rates as opposed to nominal ones. Based on expression (17) the requirement is to deflate the quantities E_t^s, W_t^s and Γ_t^s using the consumer price index as provided by the Kagera Health and Retirement Study.[44] When the quantities are deflated actual calendar time (i.e., observing a unit between rounds P_t and P_{t+1}) is no longer relevant; it only matters that the household defines *one observation* between subsequent periods t and $t+1$ if the household is observed only over two rounds (i.e., over $P1, P2$, or $P2, P3$, or $P3, P4$), *two distinct observations* between periods t and $t+1$ if the household responds over three consecutive rounds (i.e., $P1, P2, P3$ or $P2, P3, P4$), or *three distinct observations* between periods t and $t+1$ if the household is observed consecutively over the full study period (i.e., if the household responds at $P1, P2, P3$ and $P4$). However, not all observations defined in such way are valid for our analysis: an observation is valid for our analysis if children of primary school age at t remain as observations in $t+1$ so that we can then define the dynamic educational choices made for them in the household and over time. Households without school age children at t as well as households that experience the loss of a school age child between periods (i.e., between t and $t+1$) are eliminated from the sample. Table 2 shows the distribution of the total number of units observed among consecutive rounds and the distribution of these observations according to the existence of primary school age children. The same Table 2 shows that 628 households are observed consecutively between $P1$ and $P2$, 654 households are observed consecutively between $P2$ and $P3$, and 684 households are observed consecutively between $P3$ and $P4$: in total there are 1,966 units that provide information between consecutive periods ($t, t+1$). Each unit provides information regarding income growth, productivity shocks and any dynamic changes regarding educational choices for the primary school age children in the household. With regards to these children, Table 2 shows that the 1,966 units represent 5,111 children observed between consecutive time periods ($t, t+1$). The same table shows that the within sample childhood mortality rate (conditional on being age 6 to 15 at any t-period) has a negligent effect on the

[44] The KHDS provides two alternative price indices: the 'actual' CPI that reflect economic changes with regards to prices in the Kagera region, and the 'constructed' CPI that provides changes in prices reported by the survey respondents and local markets. The difference between the two indices is negligent and we choose to work with the actual CPI.

number of children for whom educational choices have to be made over two consecutive periods. In fact, only two children – both belonging to the same household – are observed at time t but die before reaching the next period $t+1$, and this happens in the first round of the survey.

Table 2: Sample dynamics by number of consecutive observations over rounds (September 1991 – January 2004)

	Observations contributing between t and $t+1$ where $t,t+1=P1,P2$	Observations contributing between t and $t+1$ where $t,t+1=P2,P3$	Observations contributing between t and $t+1$ where $t,t+1=P3,P4$	Total number of contributing observations between t and $t+1$
Total number of observations				
By number of households	834	848	852	2,534
By number of individuals (all ages)	4,896	5,424	5,968	
Total number of observations with zero primary school age children recorded in period t.				
By number of households	206	194	168	568
Respective number of individuals (i.e., ages16+)	674	705	635	
Total number of observations with the primary school age children observed between t and $t+1$				
By number of households	628	654	684	1,966
Number of adults	4,228	4,723	5,334	14,285
Number of children in age range [6,15]	1,566	1,700	1,845	5,111
Boys	795	845	900	2,540
Girls	771	855	945	2,571
Total number of observations with primary school age children at t, but have lost at least one school age child between t and $t+1$ (mortality).				
By number of households	1	--	--	1
By number of children lost between periods	2 (ages 6 and 8)	--	--	2

Note: Source, own estimates KHDS 1991-1994

Our estimates of v_t and κ_t will be based on the 5,111 children for whom a total of 1,966 units provide dynamic information on the household's productivity gains between consecutive periods conditional on educational choices made for these children. According to expression (17) each time the household decides to educate its children the investment benefits over consecutive periods equals growth rates v_t and κ_t. Alternatively, the model implication is that households observed as not sending the children to school experience only a gain of having the 'assets' children in the household equal to the rate κ_t: we may think of this rate as that associated with the intergenerational transfers described in Bell, Devarajan and Gersbach (2006). Table 3 shows the distribution of the 5,111 units (of children) according to their observed schooling regime between consecutive periods t and $t+1$.

Table 3: Distribution of 5,111 observed units of children between t and $t+1$ among alternatively schooling regimes

Schooling Regime in period t		Schooling Regime in period $t+1$	
School $e_t = 1$	Number of units 2,847 Average age: 11.9 (2.1) Median age: 12 Minimum age: 6 Maximum age: 15 Percentage boys 56.1 (0.9)	School $e_{t+1} = 1$	2,521 $(e_t = e_{t+1} = 1)$ Average age: 12.4 (2.0) Median age: 12 Minimum age: 6 Maximum age: 16 Percentage boys: 58.0 (0.98)
		No School $e_{t+1} = 0$	326 $(e_t = 1, e_{t+1} = 0)$ Average age: 12.9 (2.6) Median age: 13 Minimum age: 7 Maximum age: 16 Percentage boys: 58.1 (2.7)
No School $e_{t+1} = 0$	Number of units 2,264 Average age: 9.4 Median age: 9.4 Minimum age: 6 Maximum age: 15 Percentage boys: 46.8 (1.0)	School $e_{t+1} = 1$	379 $(e_t = 0, e_{t+1} = 1)$ Average age: 9.5 (2.0) Median age: 9.9 Median age: 9.5 Minimum age: 6 Maximum age: 16 Percentage boys: 52.4 (2.6)
		No School $e_{t+1} = 0$	1,885 $(e_t = e_{t+1} = 0)$ Average age: 9.9 (3.2) Minimum age: 6 Maximum age: 16 Percentage boys: 45.6 (1.1)

Note: Source, own estimates KHDS 1991-1994. Each of the 5,111 units defines a consecutive observation between concurrent time-periods among the 1,966 observed units as described in Table 2. These 5,111 observations are distributed among four mutually exclusive sets of dynamic educational choices (i.e., consecutively educated, consecutively non-educated and two alternative non-consecutive educational choices). For each of these four sets the table shows the weighted average age of the units and extreme ages within subgroups. All age estimates are based on weighted values using the survey sample weights; bracketed numbers show standard deviations.

As expected, Table 3 shows that when a child is sent to school at t he or she is more likely to be sent to school in the follow up period $t+1$, i.e., of the 2,847 observed attending school at t, only 326 (12% based on weighted averages) are observed as no longer at school in the months following the first period. On the other hand, a non-school status at t is more likely to be followed by a non-school status at $t+1$: 1,885 (82% based on weighted averages) of the initial 2,264 who are not at school at t remain without going to school at $t+1$. The age ranges at t are, by construction, between 6 and 15. At $t+1$ – i.e., 6 to 7 months later – the age range of ages move between 6 and 16. With respect to gender, the probability of finding a girl in the subgroup without school (53.1%) is 9.2% higher and statistically significant ($t-value: 6.9$) than the probability of finding a girl in the subgroup that goes to school (43.9%) at t. In the event that a girl does not go to school at t, her chances to attend school in periods ahead is worst than that of boys who are equally observed not attending in the first period: this is visible because the percentage of girls who are not at school in period t is 53.2%, but such percentage increases to 54.4% when we look at the subgroup $e_t = e_{t+1} = 0$. However, the percentage of boys such that $e_t = 0$ equals 46.8% and it is higher than 45.6%, the latter being the weighted percentage of boys in the subgroup $e_t = e_{t+1} = 0$: on the other hand we notice that the subgroup of boys in $e_t = 0, e_{t+1} = 1$ increases to 52.4%. In general we find that Table 3 complies with two schooling trends found in developing economies: first, it is clear that girls are less likely to attend and to continue at school and secondly, on average, primary school attendance is delayed in time: this last point is noted from Table 3 because for any subgroup the weighted mean and weighted median are almost the same suggesting that our sample is symmetrically distributed with respect to age (e.g., normally or uniformly distributed). If that is the case the mean age should be somewhere between 10 and 11 and yet the estimates show to be between 12 and 13 years of age.

The units observed such that $e_t = e_{t+1} = 1$ and $e_t = e_{t+1} = 0$ provide information from households displaying educational choices consistent with monotonic changes on the endogenous variables in the model, thus, these are households with observed behaviour that remains consistent with the assumptions underlying our modelling strategy: it is for this reason that the parameters that we estimate are based on these two subgroups from Table 3.[45] Whereas the subgroup $e_t = e_{t+1} = 0$ is sufficient to estimate κ_t for given sub-samples based on Z_t and B_t, the subgroup $e_t = e_{t+1} = 1$ is informative on the parameters v_t and κ_t, also for given sub-samples based on Z_t and B_t. We can think that the subgroup $e_t = e_{t+1} = 0$ is the counterfactual in the population to the treated (by education) subgroup $e_t = e_{t+1} = 1$. An alternative way to compare the estimates between subgroups is to think that the human capital parameter estimated from those in $e_t = e_{t+1} = 0$ are the worst case relative to the best case scenario $e_t = e_{t+1} = 1$ while all other combinations imply in intermediate degree of human capital formation. Thus, it is sufficient to estimate human capital parameters from the two subgroups $e_t = e_{t+1} = 0$ and $e_t = e_{t+1} = 1$ – the worst scenario relative to the best scenario – to provide quantitative evidence for policy advice. We also notice that we could estimate the parameter using only the household characteristics and at household level instead of allowing for each of the units (children) in the subgroups $e_t = e_{t+1} = 0$ and $e_t = e_{t+1} = 1$ to enter as independent units. However, children vary among them even if they belong to the same household, e.g., by age, school grade, gender, health-related characteristics, days spent in illness between periods, etc. It is possible that the observed differences between children in the same household may imply different

---

[45] See Lemma 1, Lemma 2 and the corresponding proofs in Appendix 1. The alternative subgroups, i.e., $e_t = (0,1), e_{t+1} = (1,0)$ each provide dynamics to the educational choices that may imply non-monotonic effects with regards to the endogenously changing variables (the value of children, human and physical assets). In principle we can estimate the human capital parameters implied by each of these two subgroups but such

contributions by different children in the same household toward the household's aggregate growth rate. It is for this reason that we choose to estimate the parameters in (17) using the 4,446 units in Table 3 as opposed to using the underlying number of households that such units represent.[46] However, for consistency with our model – monotonicity with regards to the endogenous variables in the model – we can only take units in the 4,446 whose primary school siblings in the household share a similar schooling regime: only then can we be reassured that the selected units among the 4,446 that share the same household also share the same human capital growth equation in expression (4). Based on this, a unit in the subgroup $e_t = e_{t+1} = 0$ (or likewise, in the subgroup $e_t = e_{t+1} = 1$) is valid for our analysis if all (or the larger majority) of his or her primary school age siblings in the same households have the same educational regime over time. To maximize the size of our sample we take all units in $e_t = e_{t+1} = 1$ that belongs to a household where at least 70% of the primary school age children in the household follow the same educational dynamics (i.e., $e_t = e_{t+1} = 1$). For the subgroup $e_t = e_{t+1} = 0$ we select those belonging to households 70% or more of the primary school age children in the same household show a non-schooling status over time (i.e., at most only 30% of the primary school age children in the household benefit from regular continuous school). The cut off point 70% is arbitrary and could be reduced to increase sample size. However, we believe that with positive spillage from schooling to the household occur when a large majority of the household's children attend school and this could be thought as guaranteed if more than half (e.g., 70%) of the primary school age children attend school (or no spill-over is believed to occur if 30% of less of the children are such that $e_t = e_{t+1} = 0$). Based on the 70% cut off point, only 1,636 of the initial 2,521 in the subgroup $e_t = e_{t+1} = 1$ fall into households where predominantly (+70%) most of the children benefit from schooling choices and 1,386 of the 1,885 units fall into households where predominantly (+70%) most of the children do not benefit from schooling choices. Thus, our estimates are based on a total of 3,022 units that belong to households with homogenous educational choices while the observed schooling dynamics for each unit complies with the requirements that retain the monotonic condition on the endogenously changing variables in model (4). Table 4 shows the distribution of these 3,022 units according to types $s_j(\theta)$ that take into account $j = 1, 2, 3, 4$ and $\theta = \{\theta_1, \theta_2, \theta_3\}$: the values $j = 1, 2, 3, 4$ indicate, respectively, households with both parents present ($j = 1$), with only the father present ($j = 2$), with only the mother present ($j = 3$) and with both parents absent as result of mortality ($j = 4$). On the other hand, the division $\theta = \{\theta_1, \theta_2, \theta_3\}$ defines the head of the household declared

estimates are not necessarily consistent with the assumptions underlying the model. Consequently, we cannot make similar model based conclusions for these parameters

[46] The 2,521 units with educational choices ($e_t = e_{t+1} = 1$) represent 1,531 distinct units from a possible 1,966. Likewise, the 1,885 units without educational choices ($e_t = e_{t+1} = 0$) represent 1,545 distinct units from the same set of 1,966 units (see Table 2).

educational achievement that may be zero years of education are (θ_1), at most up to 2 years in primary education (θ_2) or primary education up to the 3rd or higher level (θ_3).[47]

Table 4: Distribution of 3,022 observed valid units of children among alternatively schooling regimes and household types

		$e_t = e_{t+1} = 1$ (1,636)	$e_t = e_{t+1} = 0$ (1,386)
$s_1(\theta)$ = 1,955	$s_1(\theta_1)$	174	336
	$s_1(\theta_2)$	792	520
	$s_1(\theta_3)$	101	32
$s_2(\theta)$ = 282	$s_2(\theta_1)$	34	58
	$s_2(\theta_2)$	80	73
	$s_2(\theta_3)$	32	5
$s_3(\theta)$ = 737	$s_3(\theta_1)$	236	266
	$s_3(\theta_2)$	162	68
	$s_3(\theta_3)$	2	3
$s_4(\theta)$ = 48	$s_4(\theta_1)$	0	13
	$s_4(\theta_2)$	14	11
	$s_4(\theta_3)$	9	1

Note: Source, own estimates KHDS 1991-1994. Each of the 3,022 units belongs to households where most children receive identical educational treatment and at the same time the units display continuous educational regimes over consecutive time periods. In households s_1 both parents are present, in s_2 only the father remains present to raise to children, in s_3 only the mother remains present to raise the children and households with s_4 are households made up of orphans where a young head of households below or at age 19 is not the biological parent of the children in the household: by default these children are younger or at primary school age. In the case of s_4 both parents are absent due to death. In single-parents households the missing adult is absent permanently but not due to working away (i.e., he or she is absent due to 'never has been there' or due to mortality). Education of the household head may be zero years of education (θ_1), two years of primary education that bring basic reading and writing skills (θ_2) and beyond basic levels to higher levels of education (θ_3) that may go up to secondary education (i.e., A-levels in the Tanzanian educational system).

The definition of $s_j(\theta)$ is taken at time t with the household type remaining almost identical at $t+1$ (i.e., 6 or 7 months later). Table 4 shows that most units belong to households with both parents present (1,995) whereas for any type of household the category θ_2 is the most likely educational achievement of the household head; 1,720 units of children are associated with parents (or guardians) that have achieved some basic level of education equivalent to two years of primary school, at most. On the other hand, few children are associated with parents that have achieved some or full degree of secondary

[47] The household head is not necessarily the same as the designated household respondent although the two often coincide. The household head is defined as the person that directs the household and is the main earner. Household where the respondent is the same as the household head, he or she is below the age of 19, has never been married and has not had children but also declares that both parents are absent and dead are defined as households directed by orphans, i.e., $s_4(.)$. By sample selection criteria, if these are in our sample, these have children of primary school age and in principle, one of these could be the declared household head.

education (185 associated with θ_3) while the zero years of education occupies the second most likely type of educational level for household heads, i.e., 1,117 units have parents (or guardians) with θ_1. It is noticeable that a larger number of children with parents such that θ_1 (62% based on weighted values) are also children in an educational regime $e_t = e_{t+1} = 0$. Furthermore, Table 4 shows that for any given level of education it is significantly more likely to find single households with a female head (737) than single households with a male head (282). Fortunately, the chance that a selected unit in our analysis falls in an orphan household is less than 2% and surprisingly the distribution of units in orphan household between the two mutually exclusive educational regimes is almost the same, i.e., the regime $e_t = e_{t+1} = 1$ accounts for 49% of the orphan units while $e_t = e_{t+1} = 0$ account for the remaining 51%.

As described in Section 4, the reason why we distribute the units of interest according to the type of household $s_j(\theta)$ is because we need to apply Assumptions 10 and 11 to the data: applying these two assumptions to the data and conditional on identical e_t, e_{t+1} regime, units among the 3,022 with similar $s_j(\theta)$, similar productive characteristics Z_t and similar fertility characteristics B_t ought to experience similar growth (or depreciation) in human capital. This is the same as saying that the curvature (or gradient) of the human capital formation schedule (see expression (4) in Section 3) is the same for children that share an identical characteristics set $\Xi_t = (e_t, e_{t+1}, \mathrm{Z}_t, \mathrm{B}_t)$. Thus, any sub-cell in Table 4 has to be further distributed among cells defined by the variables in Z_t and B_t. Variables in Z_t should approximate the similarity of the productive activities among households: we take 'type of business' to define Z_t, i.e., we let Z_t contain three mutually exclusive dummy variables that explain if the household is associated with agropecuarian activities, or trading activities (e.g., shop keepers) or if the adults in the household are associated with well defined professional activities (e.g., working positions similar to those of teachers or government employees). On the other hand, variables in B_t should subdivide the cells in Table 4 so that these subdivisions contains units that belong to households with similar fertility histories. To do so we let B_t contain the variables 'age of the youngest wife present in the household' since it is the age of the youngest reproductive female that defines the households reproductive possibilities: in the event that no female spouse is present we take the age of the unique male (for s_2 type of households) or the age of the guardian (for s_4 type of households). Furthermore, the average age of the existing children may have an influence on the value that children have for the household as productive means, i.e., the average age of primary school children in the household underlines κ_t. Altogether B_t subdivides the 3,022 units in 6 different sets according to the age of the fertile female in the household ($f \leq 49$ and $f \geq 50$) and the average age of primary school children in the household that we choose to classify into three possible ranges, namely, $\eta_{\bar{g}} \in [6,9]$, $\eta_{\bar{g}} \in (9,12]$ and $\eta_{\bar{g}} \in (12,15]$. Table 5 shows how the 3,022 units are distributed according to 12 sub-cells related to $Z_t = \{d_{agro}, d_{trade}, d_{professional}\}$ and $B_t = \{d_{f\le 49,\eta(9]}, d_{f\,49\le,\eta(12]}, d_{f\,49\le,\eta(15]}, d_{f\ge 50,\eta(9]}, d_{f\ge 50,\eta(12]}, d_{f\ge 50,\eta(15]}\}$:

Table 5: Distribution of 3,022 observed valid units of children among sub-cells with common productivity and fertility histories

		$e_t = e_{t+1} = 1$ (1,636)			$e_t = e_{t+1} = 0$ (1,386)		
		TRADE $Z_t = \{d_{agro}, d_{trade}, d_{professional}\}$			TRADE $Z_t = \{d_{agro}, d_{trade}, d_{professional}\}$		
FERTILITY		d_{agro}	d_{trade}	$d_{profesional}$	d_{agro}	d_{trade}	$d_{profesional}$
$d_{f\le 49}$ 2,114	$d_{f\le 49,\eta(9]}$ 411	75	8	4	300	14	10
	$d_{f\le 49,\eta(12]}$ 1,375	731	20	20	585	7	12
	$d_{f\le 49,\eta(15]}$ 328	169	18	17	119	3	2
$d_{f\ge 50}$ 908	$d_{f\ge 50,\eta(9]}$ 138	38	0	0	96	4	0
	$d_{f\ge 50,\eta(12]}$ 550	346	12	12	172	3	5
	$d_{f\ge 50,\eta(15]}$ 220	157	9	0	52	1	1
	TOTAL	1,516	67	53	1,324	32	30

Note: Source, own estimates KHDS 1991-1994. The symbol $d_{f\le 49}$ stands for a binary indicator that equals one if the unit belongs to a household where the female associated with the head of the household is below the age of 49; likewise, the symbols $d_{\eta(9]}$, $d_{\eta(12]}$ and $d_{\eta(15]}$ identifies units in households where the average age of primary school age children is between 6 and 9, between +9 and 12 and between +12 and 15, respectively. The symbols d_{agro}, d_{trade} and $d_{professional}$ stand for agropecuarian households, households with trading and sales activities and households with a household head whose main work activity is that of a professional employee, respectively.

The number of units (and therefore households) associated with agropecuarian production far outweigh the joint number associated with sales & trade and professional activities. Likewise, a significantly large number of units belong to households with primary school age children in the rage 9 to 12 years of age. In principle our estimation procedure requires that each of the 36 cells in Table 5 is subdivided among the 24 subgroups (household's type) as given in Table 4 so that in total we should distribute the 3,022 units among 432 sub-cells.[48] All units among the 3,022 that share one of these 432 cell

[48] That is, 24 times 36 implies 864 cells but Tables 4 and 5 have a common subdivision with respect to educational regimes so that 432 of a possible 864 cells are empty by default.

are identical with regards to the position in the human capital schedule, i.e., they have equal households characteristics with respect to start up level of human capital (i.e., $s_j(\theta)$), type of productive activity (i.e., Z_t) and fertility capacity (i.e., B_t). Furthermore, each of the unique cells share identical schooling regime so that within any given cell the units are observed as either in education (so that $e_t = e_{t+1} = 1$ implies growth rates in terms of human capital equal to $v_t + \kappa_t$) or outside the educational system (so that $e_t = e_{t+1} = 0$ implies that only κ_t accounts for human capital growth rates). In practice it is clear that many of the 432 cells will be empty because of our limited sample size of 3,022. We notice that as long as a cell has a mass (even if this is only one single unit) we can identify the human capital growth rate for such cell; this is possible using expression (17) – see Section 6 below. The reason for this is that we have a sample that once weighted it is assumed to be representative of the population. In a cell has a single unit this implies that the population is characterised by such representation and, therefore, the growth rate in human capital implied by such cell is identified. The problem we face, however, is that low mass cells cannot provide sampling error information. In order to overcome the problem associated with low mass or empty cells we provide an alternative subdivision of the sample according to an alternative representation of the information set $\left\{s_j(\theta), \mathrm{Z}_t, \mathrm{B}_t\right\}$. First, we draw from Bell, Devarajan and Gersbach (2006) we assume that single parents are equally effective at transferring knowledge to their children irrespective of the parent's gender so that $s_2(\theta) = s_3(\theta)$. Second, in the developing world individuals with zero years of education are often treated with some form of literacy training that provides basic learning skills similar to those achieved with some primary education. Thus, we can group individuals in Table 4 so that θ_1 and θ_2 are assumed to be equal levels of human capital relative to those in level θ_3 . We notice that allowing for two education levels most orphans (i.e., $s_4(.)$) belong to a single subgroup; we simplify this by allowing for a third grouping that affects orphans only, i.e., once the child belongs to an orphan type of household he or she receives a similar levels of intergenerational transfers irrespective of the schooling achievement of the older sibling (household head). Based on this the assumption $s_4(\theta_1) = s_4(\theta_2) = s_4(\theta_3) = s_4$ applies. Using these three grouping alternatives we reduce the cells in Table 4 from 24 cells to 10 cells only. The cells in Table 5 can also be reduced in the following way: within households with similar θ and $s_j(\theta)$, traders and professionals are likely to be similar with regards to the intergenerational productive transfers governing the household, i.e., Z_t . This assumption increases the sample size within cells by redefining Z_t as $\mathrm{Z}_t = \left\{d_{agro}, d_{t\&p}\right\}$ where $d_{t\&p}$ is a binary indicator that equals 1 if the household's head main business is that of trader or professional. Using the new alternative sets B_t and Z_t reduces the number of cells in Table 5 from 36 to 24 cells. Overall, Table 6 shows the new distributions with 120 cells and how the 3,022 units in Table 3 are distributed among these cells according to the two educational regimes. Some of these cells remain empty so that for these the growth rate cannot be identified in the population. Furthermore, we notice that a sub-cell $e_{t,t+1}=1$ identifies $v_t+\kappa_t$ for given $\{s_j(\theta),\mathrm{Z}_t,\mathrm{B}_t\}$, but we can only disentangle κ_t from $v_t+\kappa_t$ if we can identify the counterfactual sub-cell in the population (i.e., $e_{t,t+1}=0$ for the same $\{s_j(\theta),\mathrm{Z}_t,\mathrm{B}_t\}$ subgroup in Table 6).

Table 6: Distribution of 3,022 observed valid units of children among cells with common productivity (Z), fertility histories (B) and household's type ($s_j(\theta)$)

EDUCATION REGIME 1 $e_t=e_{t+1}=0$ (1,636)	AGROPECUARIAN HOUSEHOLDS $\mathrm{Z}=\{d_{agro}=1\}$						TRADE AND PROFESSIONAL HOUSEHOLDS $\mathrm{Z}=\{d_{t\&p}=1\}$					
	$d_{f\le 49}$			$d_{f\ge 50}$			$d_{f\le 49}$			$d_{f\ge 50}$		
	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$
$s_1(\theta_1,\theta_2)$	207	435	68	40	57	18	13	7	2	4	4	1
$s_1(\theta_3)$	6	16	0	0	4	0	1	5	0	0	0	0
$s_{2,3}(\theta_1,\theta_2)$	82	122	36	56	111	34	9	7	3	0	4	1
$s_{2,3}(\theta_3)$	4	3	0	0	0	0	1	0	0	0	0	0
$s_4(\forall\theta)$	1	9	15	Not applicable	Not applicable	Not applicable	0	0	0	Not applicable	Not applicable	Not applicable

EDUCATION REGIME 2 $e_t=e_{t+1}=1$ (1,386)	AGROPECUARIAN HOUSEHOLDS $\mathrm{Z}=\{d_{agro}=1\}$						TRADE AND PROFESSIONAL HOUSEHOLDS $\mathrm{Z}=\{d_{t\&p}=1\}$					
	$d_{f\le 49}$			$d_{f\ge 50}$			$d_{f\le 49}$			$d_{f\ge 50}$		
	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$
$s_1(\theta_1,\theta_2)$	47	510	103	21	156	71	9	20	20	0	5	4
$s_1(\theta_3)$	11	46	7; κ not id	1; κ not id	16	5; κ not id	3	1	6; κ not id	0; κ not id	5; κ not id	0; κ not id
$s_{2,3}(\theta_1,\theta_2)$	17	151	52	16	166	81	0	10	4	0; κ not id	10	5
$s_{2,3}(\theta_3)$	0	8	4; κ not id	6; κ not id	8; κ not id	0; κ not id	0	6; κ not id	4; κ not id	0; κ not id	4; κ not id	0; κ not id
$s_4(\forall\theta)$	0	16	3	Not applicable	Not applicable	Not applicable	0; κ not id	3; κ not id	1; κ not id	Not applicable	Not applicable	Not applicable

Note: Source, own estimates KHDS 1991-1994. The symbol $s_1(\theta_1,\theta_2)$ explains two parents households with minimum educational achievement, $s_1(\theta_3)$ explains two parents households with primary or secondary education, $s_{2,3}(\theta_1,\theta_2)$ explains single parents households with minimum educational achievement, $s_{2,3}(\theta_3)$ explains single parent's households with primary or secondary education and $s_4(\forall\theta)$ explains orphan's households for any given level of education of the older child in the household. See footnote in Table 5 for other symbols. Cells defined as 'not applicable' implies that these do not exist in the population: by default, orphan households are defined as households where the head is a sibling aged 19 or below so that orphan households are not defined when the requirement is a productive household member age 50 or older. Cells with bold writing indicate cells with mass in both sub-samples (i.e., in both Regime 1 of education and Regime 2 of no education) so that for these cells we can identify the parameter κ_t based on the units with $e_t=e_{t+1}=0$ and use these estimates to disentangle v_t away from κ_t when estimating these for each bold cell with units $e_t = e_{t+1} = 1$. This follows the straight forward application of Assumption 10 in Section 4.

Table 6 shows that identification of the parameters κ_t and v_t (given κ_t) is possible for 29 cells whereas for the remaining cells the number of units for either of the subgroups $e_{t,t+1} = 0$ or $e_{t,t+1} = 1$ (or both) is zero and precludes identification. Besides identifying the values of κ_t and v_t for each of the 29 sub-cells, we further need to make inference on these from the sample to the population for each of these cells. We use a bootstrap procedure that re-samples randomly 1,000 from the original sample to estimate – within cell – the empirical distribution of the parameters. Based on these distributions we elicit confidence intervals and standard errors for each of the parameters associated with each of the cells. Clearly, cells in Table 6 with low mass will be estimates with a lower precision and less power than for higher mass cells.

Estimating the pertinent human capital parameters for each of the 29 cells in Table 6 requires applying the corresponding expression in (17). For example, having organized and selected the KHDS sample, Table 6 shows that there are 254 (47 and 207) units of children in agropecuarian households characterized by the presence of both parents, where the mother is a still of productive age and where the primary school age children average among them a mean age below 9. The 254 units of children are such that we observe 47 as going to school consecutively over two periods. According to our model this schooling decision on behalf of the household reflects an investment that brings an increase in human capital for the household equal to $v_t + \kappa_t$; for each of the *ith*-units in the set of 47, the two human capital parameter rates can be quantified using expression (17) so that $v_t + \kappa_t$ equals the net income growth rate over t and $t+1$ in the household (i.e., $\Delta \ln E_{i,t,t+1}$) minus the expected growth rate in earnings that is similar for households with identical productive and fertility characteristics (i.e., within cell average net growth in total earnings, $\Delta \ln W_{s_j(\theta),t,t+1} \,|\, \Xi_{t,t+1}$) and minus the household's specific net productivity gains as result of the changes in the state of the world (i.e., individual specific $\Delta \ln \hat{\pi}_{i,t,t+1}$). For each of the 47 units of children that experience a schooling decision, the net gain equals v_t *relative* to the net gain experienced by the 207 units that do not go to school but are otherwise living in identical households. That is, the 207 children in the 254 set of Table 6 such that $\Xi_{t,t+1,\Gamma} = \{\mathrm{Z}, \mathrm{B}, e_t = e_{t+1} = 0\}$ are the counterfactual set that provides an identification strategy for κ_t: this is the straightforward application of Assumption 10 that helps to disentangle v_t from κ_t for the 47 children with characteristics $\Xi_{t,t+1,E} = \{\mathrm{Z}, \mathrm{B}, e_t = e_{t+1} = 1\}$. The difference among the two groups is simply that the latter receives the treatment education. It is now clear that all we require to obtain the sample analogues of expression (17) is (i) the observed net income gains between t and $t+1$ for each of the units in Table 6 – i.e., $E_{i,t}^{s_j(\theta)}$ and $E_{i,t+1}^{s_j(\theta)}$ for the 1,636 units that live in households with a schooling choice, and the set $\Gamma_{i,t}^{s_j(\theta)}$ and $\Gamma_{i,t+1}^{s_j(\theta)}$ for the 1,386 units that did not experience a schooling choice –, (ii) an approximation of the productivity shocks $\widehat{\pi}_{it}$ and $\widehat{\pi}_{it+1}$ also for the 1,636 units that experience a schooling choice – and the approximation to $\widehat{\pi}_{it}$ for the households without a schooling choice –, and (iii) an approximation to the productive gains from non-school children implied by households that choose $e_t = e_{t+1} = 0$ as schooling regime; in the case of (iii) this productivity gain is approximated using the number of primary school children in the household that governs each of the units at t (n_{it}) and the average age of these children (i.e., $\eta_{\bar{g}it}$) with a modified version from Bell, Devarajan and Gersbach (20006) that implies a higher penalty regarding the productive capacity of children. We re-mark that the term $\Delta \ln W_{s_j(\theta)t,t+1}$ is cell specific and not unit specific thus representing the average within cell net earnings growth that is common to all units in the same cell as described in Table 6.

Income variables are readily available from the KHDS data and estimation of the productivity shocks ($\widehat{\pi}_{it}$ and $\widehat{\pi}_{it+1}$) requires the application of expression (16) to the data. Using the available information described above, Section 6 provides a detail account of the steps leading to the final estimates of the key human capital parameters as described in expression (17).

6. Results

Table 6 in Section 5 describes the subgroups for which we will estimate the key parameters explaining human capital growth rates as given in expression (17). Estimates of these rely on observed net household income over time once we have approximated the productivity shocks ($\widehat{\pi}_{it}$ and $\widehat{\pi}_{it+1}$) that are not directly observables from the data. Expressions (15) and (16) in Section 4.2 provide a latent interpretation that links these unobserved household-based productivity shocks to observed educational choices (e_t and e_{t+1}) conditional on a set of observed covariates (K_t) assumed to explain the household's preferences and characteristics that eventually lead to the within period educational choice (see Section 4.2). The link between the latent process that summarizes the within household utility from education and the productivity shock is a reservation policy π_{it}^R that explains how the household's evaluation of the world (as explained by the unobserved value $\widehat{\pi}_{it}$) determines the educational choice. Thus, a household chooses to educate the children if the state of the world is favourable, that is, if $\pi_{it} \geq \pi_{it}^R$ so that $P(e_t = 1 | \mathrm{K}_{it}; \gamma) = P(\pi_{it} \geq \pi_{it}^R | \mathrm{K}_{it}; \gamma)$. Expression (15) in Section 4 provides an interpretation of the problem using a truncated index model consistent with the requirements of a truncated distribution for the productivity shocks (see Sections 3 and 4). Expression (15) implies that we can obtain an estimate of the reservation policy as $\hat{\pi}_{it}^R = \exp(\mathrm{Z}_{it}'\hat{\gamma})$; the estimate $\hat{\gamma}$ is a vector of parameter that weights the household's preferences and characteristics (K_{it}) with regards to the educational choice (e_t) for all units in the population.

The first step in the result section is to estimate the vector $\hat{\gamma}$ based on a truncated index model as given in expression (16) where the outcome of interest is the binary educational choice e_{it} that equals 1 if the *ith*-primary school age child in the household is educated at t, and 0 otherwise. Notice that we take each primary school age child in the households to be the unit of interest, and not *the household*. This follows closely our arguments in Section 5 that aims at an efficient use of the data. Since each child is idiosyncratic within the household as result of differences in gender, age, schooling grade, health status, etc., heterogeneity among children within the household is valuable information that we take into account to understand how the household evaluates the schooling choice for each of the children at primary school age.[49] The value $\hat{\gamma}$ is estimated for K_{it} in $P(e_{it} \mid K_{it}; \gamma)$ alongside the set $[\underline{\pi}, \overline{\pi}]$ that is common for all in the population. Expression (13) – Section 4 – implies that the unobserved productivity shock received by the household and required for expression (17) – i.e., π_{it} – is in the range given by $(\pi_{it}^R, \overline{\pi})$ if and only if we observe that the choice for the child is $e_{it} = 1$. Alternatively, children that experience the choice $e_{it} = 0$ are assumed to be living in households where the unobserved productivity shock is adverse and located such that $\pi_{it} \in (\underline{\pi}, \pi_{it}^R)$. Any location in the corresponding intervals – i.e., $(\pi_{it}^R, \overline{\pi})$ or $(\underline{\pi}, \pi_{it}^R)$ – are equally valid to approximate π_{it} so that we choose to interpret π_{it} using $\hat{\pi}_{it} =$ mid-point of the interval. As noted in Table 2 – Section 5 – there are 5,111 children for whom educational choices are made at both t and $t+1$. Although we are only able to use information from 3,022 of these children – see the selection leading to Table 6, Section 5 –, all households in the population – and therefore all 5,111 children – are faced with an educational choice at each independent but consecutive point in time, i.e., all households evaluate the state of the world based on π_{it} at period t – or π_{it+1} at $t+1$ – so that we must use all 5,111 units that we observe within each period to elicit $\hat{\pi}_{it}^R$ and $[\underline{\pi}, \overline{\pi}]$ for the sample that we assume is

[49] The aim of estimating expression (16) is to obtain the value π_{it}^R for each household and the set $[\underline{\pi}, \overline{\pi}]$ that applies to all in the population. The value π_{it}^R represents the 'reservation policy' that quantifies the value of the child attending school for the household, and this value may change from child to child. Thus, we estimate the index model $P(e_{it} \mid K, \gamma)$ letting $'i'$ stand for any given child in the population aged between 6 and 15. An alternative would be to let $'i'$ be the household in the population and letting the set K include variables that 'average' the information of the primary school age children in the household (e.g., percentage males, average age, average schooling grades achieved, average number of days ill over the past month, etc). In this case the result would imply estimating a single reservation policy π_{it}^R that applies to all children in the household as if all children were equally valued with regards to attending school.

representative of the population. We notice that for any given household the state of nature changes over time as it is reflected by changes in preferences and characteristics, i.e., difference between K_{it} and K_{it+1} implies that we have to estimate two distinct truncated binary models given by $P(e_{it} \mid \mathrm{K}_{it};\gamma_t)$ and $P(e_{it+1} \mid \mathrm{K}_{it+1};\gamma_{t+1})$ thus obtaining the two time dependent sets of parameters $(\hat{\pi}^R_{it} = \exp(\mathrm{K}_{it}'\hat{\gamma}_t),[\underline{\pi},\overline{\pi}]_t)$ and $(\hat{\pi}^R_{it+1} = \exp(\mathrm{K}_{it+1}'\hat{\gamma}_{t+1}),[\underline{\pi},\overline{\pi}]_{t+1})$ as required in expression (17).[50]

Table 7 shows the choice of variables that enter $\mathrm{K}_{it+u}, u = 0,1$ divided according to five sets of household's preferences and characteristics. All estimates in Table 7 are based on sampling weights thus reflecting population values for different time periods and the educational choices underlying each of the 5,111 units. The summary statistics show that children in education are one to two years older than those in non-educational regimes. It may be that such difference in age reflects a delay in starting school – an empirical fact that is common in developing economies – and not a exclusion of school altogether. However, 75% of those in the $e_t = e_{t+1} = 0$ population are age 9 to 15, suggesting that a non-schooling status is not explained by a delay in entry (i.e., entry at ages 7 or 8 rather than the actual entry school age of 6). Sickness is equally likely among all educational regimes, but children in the non-school regime who are sick spend a significantly large periods in sickness than children in educational regimes. This may also be explained because when $e_t = e_{t+1} = 1$ children become sick the probability that these are taken to see a doctor is also statistically higher than children in alternative schooling regimes. Thus, it seems that investment on children's education is associated with investment on children's health. On average, it is more likely that boys are sent to school, i.e., the 2,561 units that experience a regime $e_t = e_{t+1} = 1$ live in households where all primary school age children have more chances to attend school than in other alternative educational regimes. Despite this, the 2,521 units live in households where – on average in the population – 66% of the household's boys within primary school age are sent to school but only 53% of the household's girls within primary school age are sent to school. In general, it seems that the data set does related to the schooling behaviour expected from developing economies.

[50] The set $[\underline{\pi},\overline{\pi}]$ are bounds on the bounded distribution that explains the common characteristics of the state of the world for all households in the same economy. The difference between households is that each receives a different – assumed – stochastic draw $\pi_{it} \sim [\underline{\pi},\overline{\pi}]_t$ that explains the household's productivity, consumption and investment choices conditional on the its contemporaneous characteristics and preferences. We assume that the unknown distribution for $\pi_{it} \sim [\underline{\pi},\overline{\pi}]_t$ is Log-Normal –see Section 4 – so that on average is rare to receive a very bad or a very good shock and with some variability most households will perceive the state of the world similarly while the household's unique characteristics implies different educational choices. In principle we could assume

Table 7: Summary statistic, preferences and characteristics (K_t and K_{t+1} based on the complete population, 5,111 units)

	$e_t = e_{t+1} = 1$ (2,521)		$e_t = e_{t+1} = 0$ (1,885)		$e_t \neq e_{t+1}$ (705)	
	K_t	K_{t+1}	K_t	K_{t+1}	K_t	K_{t+1}
Variables Unique to the child as unit						
Average age	11.8	12.4	9.4	9.9	10.6	11.02
Probability father absent(1)	0.32	0.31	0.16	0.14	0.34	0.18
Probability mother absent(1)	0.27	0.27	0.22	0.20	0.31	0.14
Probability both parents are dead	0.02	0.02	0.02	0.03	0.10	0.10
Average years spent in education	2.13	2.50	<1	<1	<1	<1
Probability child has been sick	0.41	0.43	0.24	0.24	0.22	0.28
Average number days child sick(2)	1.40	1.16	1.90	3.60	4.45	1.44
Probability child visited doctor if sick	0.16	0.16	0.04	0.06	0.22	0.22
Average BMI	16.3	16.7	16.40	16.7	16.1	16.8
Probability child has a vaccination card(3)	0.0001	0.0001	0.004	0.004	0.0003	0.0003
Probability child is a migrant(4)	0.13	Invariant	0.07	Invariant	0.30	Invariant
Probability child is pregnant if girl	0.10	0.02	0.13	0.033	0.09	0.20
Parental variable						
Probability parents are muslin	0.03	Invariant	0.09	Invariant	0.11	Invariant
Probability parents are catholic	0.59	Invariant	0.46	Invariant	0.51	Invariant
Probability parents are protestant	0.34	Invariant	0.38	Invariant	0.34	Invariant
Probability parents are other religion	0.03	Invariant	0.07	Invariant	0.04	Invariant
Average age female (spouse or head) (5)	43.7	44.2	35.6	35.9	38.7	39.23
Average probability parents were sick	0.38	0.41	0.48	0.46	0.48	0.40
Average number days parents sick(2)	6.42	4.20	4.41	3.04	4.56	1.54
Father average years in school	1.60	Invariant	1.46	Invariant	1.84	Invariant
Mother average years in school	1.10	Invariant	1.43	Invariant	1.62	Invariant
Probability trade is agropecuarian	95.5	Invariant	95.8	Invariant	96.0	Invariant
Probability trade is sales & trade	1.11	Invariant	1.72	Invariant	1.07	Invariant
Probability households are professionals	3.36	Invariant	2.47	Invariant	2.88	Invariant
Siblings related variables						
Probability PSAC works away(6)	0.33	0.25	0.33	0.31	0.36	0.34
Percentage PSA-boys in house go to school	0.66	0.63	0.36	0.36	0.43	0.48
Percentage PSA-girls in house go to school	0.53	0.54	0.30	0.28	0.39	0.40
Average age of PSAC(6)	10.91	10.7	10.61	10.85	10.51	10.70
Probability boys among PSAC	0.28	0.25	0.25	0.28	0.31	0.37
Average total kids in household up to age 15	4.83	4.90	5.24	5.31	5.11	5.16
Household related variables						
Probability house suffered dead lately(7)	0	0.0011	0	0	0.004	0
Probability house own business	0.35	0.50	0.32	0.48	0.28	0.25
Probability house gests scholarship for school	0.36	0.40	0	0	0.20	0.06
Probability household owns house	0.94	0.94	0.96	0.95	0.83	0.82
Probability house owns motorcycle or car	0.03	0.02	0.012	0.016	0.04	0.07
Probability house own TV or Radio	0.31	0.41	0.28	0.37	0.22	0.24
Probability house gets NGO help	0.70	0.88	0.68	0.80	0.64	0.68
Probability house received remittent(8)	0.77	0.85	0.76	0.81	0.83	0.60
Probability house lent money(8)	0.85	0.90	0.86	0.87	0.89	0.80
Probability house has toilet in premises	0.95	0.91	0.96	0.89	0.97	0.90
Average Household size	7.87	8.25	8.42	8.84	7.83	8.10
Average no. members visited doctor	0.52	0.50	0.52	0.53	0.49	0.48
Average BMI in household	17.95	17.98	18.07	18.13	17.99	18.12
Location (village) related variables						
Average km. distance to near school	0.020	Invariant	0.044	Invariant	0.074	Invariant
Average km. distance near health center	0.005	Invariant	0.007	Invariant	0.010	Invariant
Average km. distance to water source	0.031	Invariant	0.035	Invariant	0.056	Invariant

Note: Source, own weighted estimates KHDS 1991-1994. See Table 3 for a description of the sample with 5,111 units. 'Invariant' implies that the variable does not change between t and $t+1$. The set K_t (K_{t+1}) are characteristics and preferences at time t ($t+1$). (1) Mother or father absent due to work, illness, etc., but not dead, (2) The average takes zeros into account. (3) Household head asked to show vaccination card for the child. (4) Affects the child and households.

---

that $[\underline{\pi},\overline{\pi}]_t = [\underline{\pi},\overline{\pi}]_{t+1}$. `However, we can justify that $[\underline{\pi},\overline{\pi}]_t \neq [\underline{\pi},\overline{\pi}]_{t+1}$ because of a changing environment while in practice we can only estimate the set of parameters $[\underline{\pi},\overline{\pi}]$ with imprecision for the given sample.

(5) Refers to the youngest female not a child-relative and claims a status of head or partner-spouse of head. (6) PSAC: primary school age children, i.e., 6 to 15 years old. (7) First survey asked to declare recent deaths during past 12 months: after, the head declares any dead between t and $t+1$. (8) Takes into account households that received remittent and not expect to pay it back and (or) lent money to non-household members and not expected to get the lent money back.

Each of the variables in Table 7 enters the covariate set that aims at explaining the binary outcome 'education' (i.e., e_t and e_{t+1}) independently for each of the two time periods. Following Assumption 14 we choose a probit specification to estimate the two vectors of coefficients (γ_t and γ_{t+1}). These vectors act as weights on the preferences and characteristics (K_t and K_{t+1}) with reference to how important these variables are at explaining the education outcomes e_t and e_{t+1}. However, our probit model is a truncated probit that allows for the unobserved error term $\ln\pi$ to be bounded in the support of $[\underline{\pi},\overline{\pi}]$ thus satisfying the requirement of the value function in (5). The expression for the truncated probit specification is given in (16) and it implies estimating coefficients and bounds for each of the time period t and $t+1$, i.e., the sets $\hat{\gamma}_t,[\underline{\pi},\overline{\pi}]_t$ and $\hat{\gamma}_{t+1},[\underline{\pi},\overline{\pi}]_{t+1}$.[51] The bounds $[\underline{\pi},\overline{\pi}]_t$ and $[\underline{\pi},\overline{\pi}]_{t+1}$ are allowed to differ because the state of the world differs over time for the same population.

Table 8 shows the results from the truncated probit. Estimates for the bounds at t, i.e., $[\underline{\pi},\overline{\pi}]_t$, equal [0.04, 21], whereas $[\underline{\pi},\overline{\pi}]_{t+1}$ are estimated as [0.00001, 18]. In both cases the upper bound is not a corner solution (i.e., $\Phi(\ln\overline{\pi})<1$) so that the solution is identified for the problem at hand. When the bounds are estimated these become fixed for expression (16) and the coefficients in Table 8 are estimated accordingly. The significance of the coefficient associated with the variables in K_t (or K_{t+1}) are not of a first order concern to us: it is sufficient that a variable is regarded as part of the household's preferences and characteristics for these to enter the sets. Nevertheless, the coefficients in Table 8 provide some insights as to what determines the educational status of individuals: being an older child, the absence of the father from the household or an increasing percentage of children going enrolled in school in the household all are statistically significant at reducing the probability of education among the population represented by the 5,111 units. On the other hand, the dead of both parents (i.e., living as orphan, possibly under guardians) increases the chances of attending school. Likewise, signals that the household invests on

[51] The classic Probit corresponds to the set $[-\infty,+\infty]$ for $[\underline{\pi},\overline{\pi}]$ and would be analogous to an unbounded value function. In our case the bounds $[\underline{\pi},\overline{\pi}]$ become part of the estimation procedure as explained in expression (16), Section 4. We do this by considering a grid of values for $[\underline{\pi},\overline{\pi}]$ that cover a range between 0 and 30, i.e., our structural model requires $\underline{\pi}>0$ while for values greater than 30 the cumulative density function becomes invariant and equal to 1, i.e., $\Phi(q)=1\,\forall q>\ln(30)$. Allowing for pairs $[\underline{\pi},\overline{\pi}]$ such that $\underline{\pi}<\overline{\pi}$, we select the set ($\hat{\gamma}_t,[\underline{\pi},\overline{\pi}]_t$) that implies the highest maximum likelihood estimate.

the health of the child (e.g., higher BMI or having a medical card) increases the chances of children attending school while signals that imply lower health status in the household (i.e., households with a low average adults-BMI) are statistically significant at explaining a lower chance of attending school.

Table 8: Probit estimation. Outcome is Education ($e_t = 0{,}1$ and $e_{t+1} = 0{,}1$) conditional on preferences and characteristics (K_t or K_{t+1} , Table 7). Sample size is 5,111.

	$Period: t$	$Period: t+1$
	Coefficients (standard error)	*Coefficients (standard error)*
Variables Unique to the child as unit		
Constant	1.733 (0.545)	*2.105** (1.293)
Age	*-0.284*** (0.019)	*-0.146* ** (0.014)
Father absent	*-0.141* *(0.085)	*-0.443* ** (0.069)
Mother absent	-0.132 (0.090)	*-0.245* ** (0.072)
Both parents are dead	*0.252*** (0.121)	*0.218* ** (0.105)
Years at school so far for child	*-0.477*** (0.054)	*-0.305*** (0.033)
Number days child sick	0.068 (0.083)	-0.043 (0.043)
Body Mass Index (BMI)	*2.730*** (0.403)	*3.476* ** (0.346)
Child has a vaccination card	*0.994*** (0.412)	*1.282* ** (0.424)
Child is a migrant	0.044 (0.099)	0.170 (0.617)
Girl child is pregnant	0.063 (0.082)	-0.032 (0.078)
Parental variable		
Parents are muslin	*-0.255*** (0.138)	*-0.210*** (0.110)
Parents are catholic	*-0.333*** (0.123)	*-0.292*** (0.093)
Parents are protestant	*-0.357*** (0.132)	*-0.251* ** (0.101)
Age female (spouse or head)	0.001 (0.002)	-0.003 (0.002)
Number days parents sic	0.018 (0.020)	0.027 (0.021)
Father school years	*0.019****(0.012)	*0.043*** (0.009)
Mother school years	0.007 (0.013)	*0.016* * (0.011)
Agropecuarian household	*0.274** (0.145)	-0.004 (0.179)
Sales & trade household	*0.483* **(0.222)	0.055 (0.217)
Siblings related variables		
At least one PSAC works away	0.039 (0.059)	0.032 (0.053)
Percentage PSA-boys in house go to school	*-1.930*** (0.080)	*-1.506*** (0.064)
Percentage PSA-girls in house go to school	*-1.853*** (0.076)	*-1.482*** (0.070)
Average age of PSAC in house	*0.294*** (0.024)	*0.070*** (0.023)
Probability boys among PSAC	-0.010 (0.075)	-0.004 (0.074)
Number of kids in household ages [0,15]	0.034 (0.027)	*-0.075*** (0.021)
Household related variables		
House owns business	-0.069 (0.054)	*-0.084* * (0.053)
House owns living-house	0.182 (0.137)	*0.238* ** (0.102)
House owns motorcycle or car	-0.013 (0.151)	-0.044 (0.129)
House owns TV or Radio	*-0.095** (0.063)	*-0.162*** (0.053)
House gets help from NGO	-0.026 (0.061)	0.020 (0.059)
House received remittent	*-0.137** (0.076)	*-0.093* * (0.059)
House lent money	-0.018 (0.082)	-0.054 (0.069)
House has toilet in own premises	0.127 (0.103)	-0.080 (0.114)
Household size	0.016 (0.019)	*0.070* ** (0.014)
average in house who visited doctor at t	-0.075 (0.095)	0.002 (0.088)
Average BMI in household	*-2..68*** (0.479)	*-3.147*** (0.502)
Location (village) related variables		
Average km. distance to near school	0.005 (0.016)	*0.039* ** (0.018)
Average km. distance near health center	*-0.067*** (0.022)	*-0.126*** (0.024)
Average km. distance to water source	0.009 (0.013)	0.007 (0.013)
Truncation (1)*:*		
Lower bound $\underline{\pi}$	0.040 (0.00001)	0.00001 (0.000)
Upper bound $\overline{\pi}$	21 (0.0002)	18 (0.0004)
Diagnostics		
Maximum Likelihood value	-1,562.44	-1828.35
Pseudo-R2	0.56	0.477
LR- χ^2 *(P-value)*	3,894 (0.000)	3336 (0.000)

Note: Source, weighted estimates KHDS 1991-1994. See notes and comments from Table 3 and Table 7. The omitted category for the religion is 'other religions' and the omitted variable for household trade is 'professional type of employment'. The variables referent to time periods (e.g., time spent in sickness) and variables referent to distances (e.g., average kilometres to the nearest school) are introduced in natural logarithms. Bold and italics signals significance with two starts signalling a significance 5% or greater and one star signalling a significance at most at 10% level. The standard errors the bounds are based on a naïve bootstrap procedure that re-samples randomly with replacement f1,000 from the original data.

Allowing for Assumption (14), the coefficients $\hat{\gamma}_t$ (or $\hat{\gamma}_{t+1}$) in Table 8 weight each of the variables in K_t (or K_{t+1}) and provides the education reservation policy for the household, i.e., $\hat{\pi}_t^R = \exp(\mathrm{K}_t{}'\hat{\gamma}_t)$. Using these individual specific reservation policies (i.e., $\hat{\pi}_{it}^R$ and $\hat{\pi}_{it+1}^R$) together with the population based estimated bounds on π (i.e., $[\hat{\underline{\pi}}, \hat{\overline{\pi}}]_t$ and $[\hat{\underline{\pi}}, \hat{\overline{\pi}}]_{t+1}$), expression (13) provides the otherwise unknown individual (i.e., by household) specific productivity shocks $\breve{\pi}_{it}$ that we choose to approximate as the midpoint in the interval $\left(\hat{\pi}_{it}^R, \hat{\overline{\pi}}_t\right)$ in the case that $e_{it} = 1$ and the midpoint in the interval $\left(\underline{\pi}_t, \hat{\pi}_{it}^R\right)$ in the case that $e_{it} = 0$. Likewise, the midpoints in the intervals $\left(\hat{\pi}_{it+1}^R, \hat{\overline{\pi}}_{t+1}\right)$ and $\left(\hat{\underline{\pi}}_{t+1}, \hat{\pi}_{it+1}^R\right)$ allow for approximations for the ith unit's productivity shocks (i.e., $\breve{\pi}_{it+1}$) according to the observed educational regime also for the ith unit, i.e., $e_{it+1} = 1$ and $e_{it} = 1$, respectively. Table 9 tabulates the projected productivity shocks according to educational regime, time period and type of (parental-based) household.

Table 9: Summary statistics (means, and standard errors in brackets) for the Projected reservation values (π^R) and Projected productivity shocks ($\breve{\pi}$) by time period and type of sub-population

	$e_t = e_{t+1} = 0$ (1,885)		$e_t = e_{t+1} = 1$ (2,521)		$e_t \neq e_{t+1}$ (705)	
	$\hat{\pi}_t^R$; $\breve{\pi}_t$	$\hat{\pi}_{t+1}^R$; $\breve{\pi}_{t+1}$	$\hat{\pi}_t^R$; $\breve{\pi}_t$	$\hat{\pi}_{t+1}^R$; $\breve{\pi}_{t+1}$	$\hat{\pi}_t^R$; $\breve{\pi}_t$	$\hat{\pi}_{t+1}^R$; $\breve{\pi}_{t+1}$
$s_1(\theta_1,\theta_2)$	4.35; 2.19 (0.01); (0.004)	3.15; 1.57 (0.005); (0.002)	0.63; 10.8 (0.0013); (0.0007)	0.577; 9.29 (0.0006); (0.0003)	2.41; 5.22 (0.006); (0.010)	1.84; 6.49 (0.004); (0.0097)
$s_1(\theta_3)$	2.15; 1.10 (0.020); (0.010)	7.17; 3.58 (0.041); (0.020)	0.257; 10.63 (0.0017); (0.0009)	0.343; 9.17 (0.003); (0.0014)	0.288; 9.76 (0.0031); (0.0173)	4.74; 3.14 (0.022); (0.015)
$s_{2,3}(\theta_1,\theta_2)$	7.22; 3.63 (0.025); (0.012)	5.36; 2.68 (0.018); (0.009)	0.263; 10.63 (0.0008); (0.0004)	0.241; 9.12 (0.0006); (0.0003)	1.65; 5.24 (0.011); (0.0239)	1.39; 5.92 (0.014); (0.022)
$s_{2,3}(\theta_3)$	1.55; 0.795 (0.332); (0.167)	2.36; 1.18 (0.492); (0.246)	0.261; 10.63 (0.0082); 0.0041)	0.226; 9.11 (0.0049); (0.0024)	2.19; 6.50 (0.194); (0.358)	1.42; 5.09 (0.209); (0.341)
$s_4(\forall\theta)$	6.27; 3.15 (0.11); (0.055)	6.53; 3.27 ((0.093); (0.047)	0.322; 10.66 (0.0085); (0.0043)	0.345 ; 9.17 (0.00749; (0.0037)	1.27; 8.90 (0.0917); (0.131)	2.58; 3.21 (0.138); (0.147)

Note: Source, own estimates KHDS 1991-1994. See footnotes in Table 6. Standard errors are based on the asymptotic standard errors of the mean for any given sub-cell in the table. All values are estimated using the sample weights provided in the KHDS 1991 -1 994 survey.

The reservation values (or productivity shocks) cannot be interpreted on their own. Instead, these should be interpreted in comparison between different household types, according to educational regime and according to dynamic changes over time. We recall that for any individual or household in the population, at time t, the reservation value $\hat{\pi}_t^R$ summarizes the relative effect of preferences and characteristics with regards to educational choices. Our modelling strategy implies that a higher value of $\hat{\pi}_t^R$ lowers the possibility of entry into the education system; that is, schooling at t (i.e., $e_t = 1$) is associated with the event that $\pi_{it} > \pi_{it}^R$ suggesting that the household evaluates the state of the world (i.e., π_{it}) as sufficiently high as to choose education relative to the reservation value placed on education by the household (i.e., π_{it}^R). Given that π_{it} is distributed so to represent the population's view of the state of nature, the upper tail in the distribution of π_{it} will be thin, i.e., it is less likely to receive a high productivity shock (a very high and positive shock) so that $\pi_{it} > \pi_{it}^R$ is true. Therefore, household holding high reservation values on education are less likely to send their primary school age children to school in the same way as an unemployed individual is less likely to enter into paid labour market activities if she holds a very high (subjective) reservation wage.[52] Table 9 shows that our data complies with the underlying assumptions in the model, i.e., individuals in the category $e_t = e_{t+1} = 0$ belong to households whose preferences and characteristics leads to high value for $\hat{\pi}_t^R$ relative to the values estimated for households with children such that $e_t = e_{t+1} = 1$. The consistency of these estimates with the assumptions implied by the modelling strategy simply suggest that the variables in K_t (and K_{t+1}) are correctly capturing those characteristics that are important at defining the educational choices for the children in the household. Moreover, comparing the dynamic change of $\hat{\pi}_t^R$ within educational regime (i.e., within sub-groups $e_t = e_{t+1} = 0$ and $e_t = e_{t+1} = 1$ but as these change over time) we see that for most of the $s_j(\theta)$ -type the changes between $\hat{\pi}_{t+1}^R$ and $\hat{\pi}_t^R$ are also consistent with our modelling assumptions. That is, our modelling assumption implies that when a child is sent to school at t (i.e., $e_t = 1$) the *effort* of sending the child to school at $e_{t+1} = 1$ is lower. This should be reflected so that for $e_t = e_{t+1} = 1$ we find that $\hat{\pi}_t^R \geq \hat{\pi}_{t+1}^R$,

[52] The interpretation is analogous to that of voluntary choosing between work and unemployment in countries with well developed labour markets. In developed economies with a sufficient vacancy flow the effect of unemployment insurance is that of making fixed cost of entry into the labour market as the key factor that determines the difference between (voluntary) unemployment and employment. Fixed cost of entry are items such as family characteristics such as number of kids, household's preferences such as taste for leisure, etc. Such preferences and characteristics determine a reservation wage w_t^R that may vary over time but may also differ according to characteristics (e.g., the reservation wage is higher for a female with young children than a similar aged female without children). Entry into paid activities occurs if $w_t > w_t^R$, where the value w_t summarizes the state of the labour market (i.e., the state of the world). For example, w_t can be the minimum wage offer and w_t^R is the ongoing unemployment benefits that varies by individual and over time. In the labour market literature the value w_t^R is unobserved for the unemployed and preference and characteristics are often used to elicit such values in the population. Likewise, in our specific problem, the values π_t and π_t^R are not observed and we use preferences and characteristics to elicit these from the data.

and alternatively, for $e_t = e_{t+1} = 0$ it should be the case that $\hat{\pi}_t^R \leq \hat{\pi}_{t+1}^R$, i.e., choosing no school at t becomes more of an effort to send the child to school at $t+1$.[53] Table 9 shows that on average, for the non-schooling regime (i.e., $e_t = e_{t+1} = 0$) the expected dynamic relation between the reservation values is found for all household except for types $s_1(\theta_1, \theta_2)$ and $s_{2,3}(\theta_1, \theta_2)$. On the other hand, we find that for the schooling regime sub-sample (i.e., $e_t = e_{t+1} = 1$) the within $s_j(\theta)$-type mean difference in reservation values ($\hat{\pi}_t^R - \hat{\pi}_{t+1}^R$) is either statistically insignificant or always positive, i.e., in period $t+1$ the reservation value of sending the primary school age children is lower or equal if these were initially sent to school in period t: another way to read this is that the value of this children as workers o idle (the alternatives to a schooling regime) has decreased as result of the initial schooling investment in period t. Finally, Table 9 provides summary statistics for the residual category with non-monotonic educational choices. The estimates for the reservation values do not necessarily comply with the assumptions in our model while comparison between $\hat{\pi}_{t+1}^R$ and $\hat{\pi}_t^R$ suggest that the sub-population have preferences for schooling that places them between the two extreme sub-groups, i.e., between $e_t = e_{t+1} = 0$ and $e_t = e_{t+1} = 1$. Our analysis below is based on the selection of individuals from the sub-groups $e_t = e_{t+1} = 0$ and $e_t = e_{t+1} = 1$ that comply with the selection criteria implied in Table 6.

At this point we are ready to implement expression (17) that leads to estimates of the final policy parameters central to the structural model in (4). For this purpose, and together with knowledge of household type $(s_j(\theta), \mathrm{B}, \mathrm{Z})$, projected productivity shock $(\breve{\pi})$, and the approximation to children productivity in non-schooling households ($\Upsilon = W\sqrt{n} \cdot \eta_{\bar{g}}$), we require real net household income growth rate between periods t and $t+1$. The KHDS 1991-1994 data provides two distinct measures of income both of which are equally valid and differ only with regards to how business related profits vary according to answered questions: *income type 1* makes use of declared business costs and declared business revenue to estimate net returns from business wherease *income type 2* makes use of direct responses to questions on net profits to estimate total household income over the periods.[54] Table 10 shows summary statistics for the two measures of income allowing for the the samples as described in Table 6 since these are the sub-groups finally employed to estimate the parameters in expression (17).

---

[53] As this was previously suggested, the relative cost can be measured by comparing the opportunity costs implied by sending a child to work if the child had previously participated at the education system and, likewise, the opportunity cost lost in terms of non-learning production gains of non-school children.

[54] When families are asked the first time, total household income is measured over the last 12 calendar months, but for subsequent periods incomes refer to that over a 6 months period. We convert all data to reflect 6 months of net real household income.

Table 10: Summary statistics (means, and standard errors in brackets) for net real household income. Selected sample based on monotonic educational regimes and homogenous educational choices in households

	$e_t = e_{t+1} = 0$ (1,386)		$e_t = e_{t+1} = 1$ (1,636)	
Income Type 1	t	$t+1$	t	$t+1$
$s_1(\theta_1,\theta_2)$	169,886 (576)	149,885 (379)	121,305 (260)	133,878 (416)
$s_1(\theta_3)$	143,727 (3,079)	66,247 (1,187)	249,183 (955)	242,501 (1,114)
$s_{2,3}(\theta_1,\theta_2)$	90,680 (277)	95,549 (304)	68,940 (122)	58,432 (127)
$s_{2,3}(\theta_3)$	230,377 (15,290)	249,488 (31,360)	200,926 (4,527)	181,271 (4,314)
$s_4(\forall\theta)$	107,407 (1,372)	73,250 (1,087)	48,519 (797)	46,356 (592)
Income Type 2				
$s_1(\theta_1,\theta_2)$	173,228 (591)	158,607 (409)	120,732 (265)	153,631 (443)
$s_1(\theta_3)$	110,239 (1,792)	66,387 (1,190)	241,090 (951)	250,186 (1,038)
$s_{2,3}(\theta_1,\theta_2)$	90,101 (283)	95,475 (301)	69,127 (126)	59,468 (130)
$s_{2,3}(\theta_3)$	230,587 (15,330)	251,990 (32,607)	178,210 (4,962)	184,816 (4,185)
$s_4(\forall\theta)$	107,565 (1,371)	73,516 (1,092)	53,438 (803)	55,122 (504)

Note: Source, own estimates KHDS 1991-1994. See footnotes in Table 6 and the same table for an explanation of the sample selection criteria. INCOME TYPE 1 refers to net income when business income is estimated taking away costs from gross revenue. INCOME TYPE 2 is identical to TYPE 1 but now net income is estimated using direct responses to questions on net revenue (profit). Both incomes are net and in real terms with home production that has entered using local market prices. Standard errors are based on the asymptotic standard errors of the mean for any given sub-cell in the table. All values are estimated using the sample weights provided in the KHDS 1991 -1 994 survey. All values reflect real net income in Tanzanian shillings with base year and month in September 1991. The USA conversion rate for TZS in September 1991 was approximately 219.2 TZS per 1$USA.

Table 10 shows that on average the two income measures lead to similar with $s_j(\theta)$ averages so we choose to comment and employ *income type 2* that relies on less number of recall questions. Based on *income type 2,* we see that average income in similar $s_j(\theta)$ -types are not necessarily always higher when $e_t = e_{t+1} = 1$ is in fact the educational regime.[55] In fact, for both types of income, income is on average higher for units in $e_t = e_{t+1} = 1$ when the households are two parents household with higher levels of education (i.e., for $s_1(\theta_3)$ -type). What is clear from the data is that units in the sub-group $e_t = e_{t+1} = 0$ belong to households such that *real* income can decrease significantly the short interval of time implied between t and $t+1$ (i.e., over a 6 months period). On the other hand, units in $e_t = e_{t+1} = 1$ belong to households that on average will experience income growth over time (except for single parent households with lower educational achievements, i.e., $s_{2,3}(\theta_1,\theta_2)$). In both cases (i.e., both for sub-groups $e_t = e_{t+1} = 0$ and $e_t = e_{t+1} = 1$) the larger majority of households (96% and 93% respectively) belong to the

[55] It is important to notice that net income includes scholarship grants for schooling in the event that a household has reveived these over the period.

agropecuarian industry for matters of production. Thus, crop variability cannot explain the difference in income dynamics between the two educational regimes. Instead, the difference may reflect differentials in the actual position within the industry (e.g., labourer as opposed to owner of business), and therefore such income difference within similar household types already picks up some of the household's differential in productivity gains and capacity over time. Table 10 also shows that in terms of average income single households with low education (i.e., $s_{2,3}(\theta_1,\theta_2)$ are closely related to orphan households (i.e., $s_4(\forall\theta)$)where an elder sibling acts as a guardian. Nevertheless, at levels and over time, and for any given income measure, on averge, orphans are the lower income earners in this society.

We allow for *income type 2* to be the basis for our estimates of the parameters in expression (17). The variable $E_{it}^{s_j(\theta)}$ refers to the observed ith household net real income for anyone in $s_j(\theta)$ at time t in the event that the unit belongs to the sub-group $e_t = e_{t+1} = 1$. Likewise we define $E_{it+1}^{s_j(\theta)}$ at time $t+1$. Similarly, the terms $\Gamma_{it}^{s_j(\theta)}$ and $\Gamma_{it+1}^{s_j(\theta)}$ indicate the income values of any given ith household in $s_j(\theta)$ in the event that the unit is in the sub-group $e_t = e_{t+1} = 0$. The human capital growth parameters v_t and κ_t are estimated separately for each of the cells in Table 6. Based on our model, all units sharing the same cell in Table 6 belong to households with identical levels of human capital and identical potential to change in terms of human capital. Thus, following a similar assumption as that implied in Hekman, Taber an Lochner (1998) each of these cells is characterized by an underlying average income that explains the 'skill-type' of the households. Let this value be $W_t^{s_j(\theta)}$ (and $W_{t+1}^{s_j(\theta)}$) common to all within each of the cells in Table 6: we use within cells averages in net real earnings to approximate the two values $W_t^{s_j(\theta)}$ and $W_{t+1}^{s_j(\theta)}$ for each of the cells. Thus, following expression (12) and it derivative in (17), for any individual unit in Table 6, the household's growth earning rate experienced over time (i.e., either $\Delta\ln E_{it,t+1}^{s_j(\theta)} = \ln E_{it+1}^{s_j(\theta)} - \ln E_{it+1}^{s_j(\theta)}$ or $\Delta\ln\Gamma_{it,t+1}^{s_j(\theta)} = \ln\Gamma_{it+1}^{s_j(\theta)} - \ln\Gamma_{it+1}^{s_j(\theta)}$) is the sum of three components: *(i)* the average gains exclusive to those with identical skills in the production process (i.e., $\Delta\ln W_{t,t+1}^{s_j(\theta)} = \ln W_{t+1}^{s_j(\theta)} - \ln W_{t+1}^{s_j(\theta)}$), *(ii)* the average gain – or loss – that may result from the households and unit specific idiosyncratic changes that affects the educational choice (i.e., $\Delta\ln\breve{\pi}_{it,t+1} = \ln\breve{\pi}_{it+1} - \ln\breve{\pi}_{it}$) and *(iii)* the gains that children bring to the household as result of their existence as potential productive units for the houseold. The gains implied by *(iii)* vary according to the educational regime of the child. Relative to a state of non-schooling, a unit that is observed to benefit in full from school – i.e., units in $e_t = e_{t+1} = 1$ – is a non-productive unit whose schooling status implies an investment gain to the household measured as v_t : this is the human capital gain to the household of investing in education. But children also learn from belonging to a given household type; we measure this as the household's intergenerational transfer represented by κ_t in expression (12) and (17). The estimate κ_t is analogous to that in Bell, Devarajan and Gersbach(2003) but now we provide a short-run micro-econometric interpretation. Finally, units in the sub-group $e_t = e_{t+1} = 0$ are not sent to school and will most likely be engaged in household or non-household productive activities that bring the households an amount that can be approximated as $\Delta \ln_{it,t+1}\left(W_t^{s_j(\theta)} \cdot \left(n^{1/2}\eta_{\bar{g}}\right)\right)$, where n stands for the number of primary school children in the household, $\eta_{\bar{g}}$ depends on the average age of the primary school age children in the household and helps the approximate the productivity of a child relative to that of an adult (see the proposal for Regime 2 in Section 4.1, and Bell, Devarajan and Gersbach (2003) for an alternative but similar specification regarding the productivity of primary school age children).

Allowing for data on income as explained in Table 10 (*income type 2*) and using the estimated productivity shocks as explained in Table 9 (see Table 7 to Table 9), we are now ready to estimate the human capital rates experienced by each of the sub-populations implied in Table 6. These human capital rates are identified and equal to κ_t for all non-empty cells in the sub-group $e_t = e_{t+1} = 0$, thus, all cells that register zero units in the population are defined as 'not identified' cells. We emphasize that this is not a data or sample selection problem; our sample is in fact representative of the Kagera region so that empty cells are in fact household types that are not represented in such population. In the case of the sub-group $e_t = e_{t+1} = 1$ the assumption is that the units will experience human capital changes resulting from intergenerational transfers (κ_t) as well as from the effects of education (v_t). Thus, for any given cell in the sub-group $e_t = e_{t+1} = 1$, changes in human capital resulting from the existence of primary school children in the household can be identified if and only if the data also identifies the analogous cell with respect to the information set $\{s_j(\theta), \mathrm{B}, \mathrm{Z}\}$ in the sub-group $e_t = e_{t+1} = 0$ (see Assumption 10), i.e., learning as result of intergenerational transfers (κ_t) is identical for households with identical productivity characteristics $\{s_j(\theta), \mathrm{B}, \mathrm{Z}\}$ even if these do not share the same educational regime. The estimated parameters (κ_t and v_t) define the gradients for $\Delta \ln \Lambda_{t,t+1}$, i.e., $\Delta \ln \Lambda_{t,t+1} = v_t \breve{\pi} + \kappa_t$ for the sub-group $e_t = e_{t+1} = 1$ and $\Delta \ln \Lambda_{t,t+1} = \kappa_t$ for the sub-group $e_t = e_{t+1} = 0$. The top 5 rows in Table 11 show cell specific estimates of the human capital growth rate purely from intergenerational transfers, i.e., κ_t, for the non-school educational regime ($e_t = e_{t+1} = 0$). The bottom 5 rows in Table 11 show the estimates of human capital growth rate that result from units that receive the benefits from their household's investments on their education, i.e., the addition of $(v_t \cdot \breve{\pi}_{cell-specific} + \kappa)$, where $\breve{\pi}_{cell-specific} = \left(1/_{i \in cell}\right) \sum_{i \in cell} \breve{\pi}_i$, i.e., we assume that each unit perceives the same stochastic shock than similar units in the cell. Together with the estimated human capital growth rates, each cell shows the number of units employed in the estimation of the actual value (see Table 6). Estimated parameters that are statistically significant are shows in bold and italics.

Table 11: Estimates of human capital growth rates (κ_t and v_t) by household types according to productivity class and educational regimes

EDUCATION REGIME 1 $e_t = e_{t+1} = 0$ (1,386) κ	AGROPECUARIAN HOUSEHOLDS $Z = \{d_{agro} = 1\}$						TRADE AND PROFESSIONAL HOUSEHOLDS $Z = \{d_{t\&p} = 1\}$					
	$d_{f\le 49}$			$d_{f\ge 50}$			$d_{f\le 49}$			$d_{f\ge 50}$		
	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$
$s_1(\theta_1,\theta_2)$	***-0.024*** 207	***0.174*** 435	***0.396*** 68	***0.026*** 40	***0.455*** 57	0.545 18	***0.968*** 13	0.222 7	0.437 2	-0.015 4	0.017 4	0.426 1
$s_1(\theta_3)$	-8.189 6	-0.043 16	Not identified	Not identified	0.254 4	Not identified	-1.239 1	-0.106 5	Not identified	Not identified	Not identified	Not identified
$s_{2,3}(\theta_1,\theta_2)$	***-0.009*** 82	***0.003*** 122	***0.265*** 36	***0.216*** 56	***0.123*** 111	***0.574*** 34	0.215 9	0.036 7	-0.170 3	Not identified	0.185 4	-0.004 1
$s_{2,3}(\theta_3)$	-0.479 4	-0.064 3	Not identified	Not identified	Not identified	Not identified	0.336 1	Not identified	Not identified	Not identified	Not identified	Not identified
$s_4(\forall\theta)$	0.173 1	-0.357 9	0.378 15	Not applicable	Not applicable	Not applicable	Not identified	Not identified	Not identified	Not applicable	Not applicable	Not applicable
EDUCATION REGIME 2 $e_t = e_{t+1} = 1$ (1,636) $v\breve{\pi} + \kappa$	AGROPECUARIAN HOUSEHOLDS $Z = \{d_{agro} = 1\}$						TRADE AND PROFESSIONAL HOUSEHOLDS $Z = \{d_{t\&p} = 1\}$					
	$d_{f\le 49}$			$d_{f\ge 50}$			$d_{f\le 49}$			$d_{f\ge 50}$		
	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$	$d_{\eta(9]}$	$d_{\eta(12]}$	$d_{\eta(15]}$
$s_1(\theta_1,\theta_2)$	***-0.209*** 47	***0.178*** 510	***0.233*** 103	***0.179*** 21	***0.426*** 156	***-0.210*** 71	-0.549 9	***0.336*** 20	***0.081*** 20	-- 0	0.017 5	0.137 4
$s_1(\theta_3)$	-0.023 11	***0.076*** 46	7; κ not id	1; κ not id	***0.086*** 16	5; κ not id	0.169 3	0.169 1	6; κ not id	0; κ not id	5; κ not id	0; κ not id
$s_{2,3}(\theta_1,\theta_2)$	***0.132*** 17	***-0.026*** 151	***0.212*** 52	***0.162*** 16	***0.094*** 166	***0.086*** 81	-- 0	-0.067 10	0.125 4	0; κ not id	0.395 10	0.576 5
$s_{2,3}(\theta_3)$	-- 0	0.077 8	4; κ not id	6; κ not id	8; κ not id	0; κ not id	-- 0	6; κ not id	4; κ not id	0; κ not id	4; κ not id	0; κ not id
$s_4(\forall\theta)$	-- 0	***0.159*** 16	0.152 3	Not applicable	Not applicable	Not applicable	0; κ not id	3; κ not id	1; κ not id	Not applicable	Not applicable	Not applicable

Note: Source, own estimates KHDS 1991-1994. See comments and notes in Table 6. Each cell shows the estimated value for human capital growth (i.e., κ for the first 5 top rows and v for the 5 last rows). The numbers in brackets reproduce the number of observations in Tabl6. Sampling variance is based on a naïve bootstrap proceedure that resamples randomly with replacement from the original data: using a two tail test, bold and italics indicates significance at least at a 10% level.

Table 11 shows that estimated parameters can be both negative and positive. A negative parameter implies depriciating human capital whereas a positive parameter implies increases in human capital between consecutive time periods. Thus, given $\Lambda_{t+1}/\Lambda_t = \exp(\mu_t)$, $\mu_t = \{\kappa_t, v_t \cdot \breve{\pi}_t + \kappa_t\}$, a negative value for μ_t would imply that relative to Λ_t the value Λ_{t+1} (i.e., next periods human capital or productive capacity) has decreased, whereas if μ_t is positive the implication is that relative to Λ_t the value Λ_{t+1} has increased. Only cells with statistically significant values are analysed since these are the cells that allow us to infer from the sample to the population: for all other cells – mostly for cells with low density – the samping variance or lack of it renders the cell statistically insignificant. Overall the evidence in Table 11 seems to suggest that on average intergenerational transfers remain the strongest component in terms of human capital formation. For example, $s_1(\theta_1,\theta_2)$ agropecuarian households with young fertile mothers ($d_{f\leq 49}$) and primary school children aged between 10 and 12 experience a 17.4% human capital growth if the children are not sent to school and 17.8% if these children are sent to school. Thus, for this particular sub-group in the population the effect for the household of educating is positive but relatively small when compared to the effect that intergenerational transfers have on human capital. An alternative way to examine the strength of intergenerational transfers is to compare growth rates for identical $\{\mathrm{B}_t, \mathrm{Z}_t\}$ households that differ according to $s_j(\theta)$ but have similar schooling regimes. For example, we can compare '$s_1(\theta_1,\theta_2)$-type' agains '$s_{2,3}(\theta_1,\theta_2)$-type' both holding identical $\{\mathrm{B}_t = d_{f\leq 49}, d_{\eta(12]}, \mathrm{Z}_t = d_{agro}\}$ and under schooling regime $e_t = e_{t+1} = 0$. In this case households differ only with regards to parental support: we find that when both parents are present intergenerational transfers imply an increase in human capital equal to 17.4%, whereas when only one parent is present – i.e., $s_{2,3}(\theta_1,\theta_2)$ – the human capital intergenerational transfer equals only 0.3%, a value that is relatively close to the concept of zero intergenerational transfers. Table 11 shows that young agropecuarian households (i.e., $d_{\eta(9]}$ and $d_{f\leq 49}$) with a non-schooling regime (i.e., $e_t = e_{t+1} = 0$) experiences significant human capital depreciation for types $s_1(\theta_1,\theta_2)$ and $s_{2,3}(\theta_1,\theta_2)$; despite its significance in both cases the depreciations are low in magnitude (2.4% and less than 1%, respectively). On the other hand, human capital depreciation is not exclusive to those in a non-educational regime. Table 11 shows that some household types in the sub-group $e_t = e_{t+1} = 1$ are characterised by statistically significant depreciating human capital: in particular, our data suggest that low educated two parent's households in agropecuarian production (i.e., $s_1(\theta_1,\theta_2)$) experience a 20.9% drop in human capital when the children sent to school are relatively young (i.e., ages between 6 and 9). Likewise, depreciaton is associated with low educated single parent households in agropecuarian production with primary school age children in the age bracket 10 to 12; in their case the estimate is statistically significant but low in magnitude with a drop in human capital estimated at a 2.6%.

The estimates in Table 11 can be graphically displayed allowing for the generic expression in human capital as defined in expression (4). In the case of a non-schooling regime, i.e., $e_t = e_{t+1} = 0$, the expression is $\Lambda_{t+1}/\Lambda_t = \exp(\kappa_t)$, whereas for $e_t = e_{t+1} = 1$ this becomes $\Lambda_{t+1}/\Lambda_t = \exp(\kappa_t, v_t \cdot \breve{\pi}_t)$. In order to understand the effect of a household type (i.e., $s_j(\theta)$) on human capital and the difference according to the two alternative schooling regimes, we can interpret each set of column cells for any given $s_j(\theta)$ as the change on human capital over time. For example, agropecuarian households of type $s_1(\theta_1,\theta_2)$ with children in age range 6 to 9 (i.e., $d_{\eta(9]}$) and younger mothers (i.e., $d_{f\leq 49}$) cannot provide information on human capital growth rates for periods ahead when their children would be in the next age-cohort (i.e., $d_{\eta(12]}$). Nevertheless, such information is available for similar household types with children in the cohort $d_{\eta(12]}$. Thus, allowing cohort changes to define the dynamic changes of human capital over time we have an interpretation of the dynamics of human capital over time for each of the different $s_j(\theta)$ households in Table 11. Let the cohort $d_{\eta(9]}$ in households $d_{f\leq 49}$ to be interpreted as the first period of any given household, that is, $\Lambda_{t+1}/\Lambda_t = \Lambda_1/\Lambda_0$ for households with younger children (i.e, $d_{\eta(9]}$ and $d_{f\leq 49}$). In the case of $s_1(\theta_1,\theta_2)$ agropecuarian households in a non-schooling regime we have that $\Lambda_1 = 0.976\cdot\Lambda_0$, i.e., next periods productive capacity in such household is 97.6% relative to contemporaneous human capital. Normalizing $\Lambda_0 = 1$ we can build the full sequence of human capital relative to this normalizing constant. Thus, following the same example we have that $\Lambda_1 = 0.976|_{\Lambda_0=1}$, $\Lambda_2 = 1.19\cdot\Lambda_1$ so that in fact we can express $\Lambda_2 = 1.19\cdot\Lambda_1 = 1.19\times 0.976|_{\Lambda_0=1} = 1.61|_{\Lambda_0=1}$. Figure 1 and Figure 2 provide the dynamic plots that interpret the evolution of human capital for agropecuarian households $s_1(\theta_1,\theta_2)$ and $s_{2,3}(\theta_1,\theta_2)$; it is for these two household types that the data is able to identify the required sequence of eight parameters for either one of the schooling regimes. For completeness of analysis the same two figures plot the information available for orphan households, i.e., $s_4(\theta)$.[56] Figure 1 shows the effects on human capital of a non-schooling regime and Figure 2 shows the same in the case of a schooling regime; in both cases with reference to the schooling choise that parents make for the primary school age children in the household. Given that 96% of the units in our data belong to agropecuarian households we restrict the graphical interpretation to this production type.[57]

[56] Orphans are those who experience the lower rates of intergenerational. Clearly, our data can identify some of these but its significance is questionable as result of the low sample size. Nevertheless, we use all available information to understand the relative difference between three household types that cover all spectrum of household's composition in the Kagera region. For completeness we have assumed that human capital in Period 1 is such that $\Lambda_1 = \Lambda_0$ for orphans in sub-group $e_t = e_{t+1} = 1$.

[57] Clearly, not all 'cell by cell' estimates in Table 11 are significant. However, when identified we use the estimates to allow for a continous graphical interpretation.

Figure 1: Human Capital, Agropecuarian households with non-schooling regimes

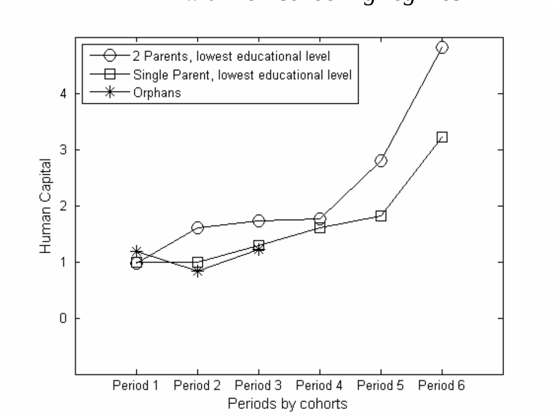

Figure 2: Human Capital, Agropecuarian households with schooling regimes

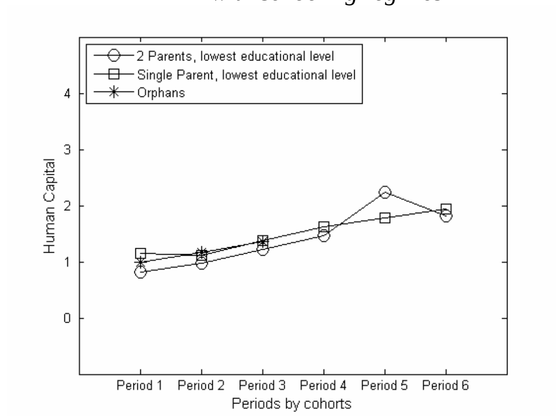

Figure 1 (a non-schooling regime) shows that for any given cohort (i.e., at any age implied by the primary school age children in the households) and relative to single parents households, two parent households are consistently higher providers of intergenerational transfers of human capital to their children. At the beginning of the children's live as productive units (i.e., in Period 1 and when children are between the ages of 6 and 9) the difference in the productive capacity between two parents and single parent households is negligent. Over time the intergenerational transfers between the two types of households widens significantly to become 33% higher for children in two parent households relative to those in single parent households. Figure 2 refers to units under schooling regimes: this figure shows that in the case of a schooling regime, the household type (i.e., comparing single parent to two parent households) makes no difference with regards to the inherited productive capacity that the child receives: with regards to human capital transfers there is a negligent difference at either levels (the value of human capital) or in terms of growth (the gradient implied by the plot). As far as data allows (Periods 1 to 3) and relative to the two other household types considered, orphans also experience similar trends in human capital as result of attending school.

Altogether Figure 2 would seen to suggest that schooling can make up for the effect of household differences, i.e., attending school helps to close the human capital gap between children in disadvantaged households (i.e., single parent and orphan households) and children in potentially less adversed environments (i.e., children living in two parent households). However, when we compare Figure 1 to Figure 2 we see that – at least for Period 1 to Period 4 – the effect of schooling on human capital is negligent: inspecting the two figures over these four periods suggests that estimates of intergenerational transfers (i.e., $\hat{\kappa}_t$ in Figure 1) drives the human capital gradient so that when schooling effects are added (i.e., $\hat{v}_t\breve{\pi}+\hat{\kappa}_t$ in Figure 2) the gradient remains relatively stable over time. More specifically, in two parent households the increase in human capital between Period 1 and Period 4 equal to 81.5% and 80.4% for non-schooling and schooling regimes, respectively. Both regimes start up with similar human capital (at levels) in Period 1 so that for two parent households we cannot detect the effects of education on human capital. In the case of single parent households the increases in human capital over the same four periods are 62.2% for the non-schooling regime and 41.6% for the schooling regime. The only difference by schooling regime between single parent households is that of levels in Period 1: at such point human capital at levels is 14.2% higher for children in the schooling regime relative to the level of human capital for those in a non-schooling regime. However, a slightly growth spread over the four periods for those in a non-schooling regime implies that the difference at levels has drop from 14.2% to only 0.5% once the households reach Period 4. Beyond Period 4 (i.e., for cohorts of primary school age children between the ages of 13 and 15) the effect of schooling becomes adverse on human capital, that is, all households in both types of schooling regimes are associated with increasing levels of human capital. Howerve, a non-school regime can experience growth rates in human capital at household level that are significantly higher than those experienced by units in a schooling regime. For example, comparing Figure 1 to Figure 2 shows that in Period 6 the two parent households in the non-school regime attain a level of human capital that is 63% higher than the two parent households in the schooling regime. Likewise, non-schooling single parent households attain levels of human capital in Period 6 that are 40% higher than single parent households in the schooling regime.

Overall, the evidence analysed so far suggest that intergenerational transfers (i.e., differentials in terms of parental presense) is far more important than the schooling regime in oder to determine the growh rate in terms of human capital in the household. This reasoning would be in line with the findings that affect orphans households thus far not analysed. Such rates apply only to Periods 1 to 3. We notice from Figure 1 that a non-schooling regime leads to human capital growth after an initial period of depreciation. This would suggest that when parents are not present (i.e., an older sibling acts as a parent) there is no productivity gains from very young working children (i.e., Period 1) but that as these get older some intergenerational transfers occurs thus leading to some gains in terms of human capital. The gains, however, are very small, i.e., Figure 1 shows that comparing Period 2 to Period 3 human capital has increased by 45.9% after an initial period of depreciation that reduced human capital by 35.7%, thus the net gain in human capital is 10.2%. On the other hand orphans that attend school experience a 16.4% net increase in human capital over the three periods. This suggests that when much of the human capital gains cannot be justified by the presence of parents, schooling makes a positive and significant difference with regards to the productive capacity of children in developing economies.

7. Conclusions

This paper provides a structural framework to theoretically and empirically analyse endogenous human capital formation at the household level and in the settings that characterises a developing economy. The basic structural framework determines the dynamic evolution of human capital – i.e., the productive capacity within the household – in the presence of a stochastically changing environment often characterised by adverse health conditions. The stochastic state of the world that conditions the household’s productive capacity often determines the household's schooling investment choices for its primary school age children. The structural model assumes two distinct educational regimes for the household, namely, a school regime where school age children attend school regularly, or a non-school regime where school age children do not attend school and are therefore assumed to produce consumption goods for the household: the choice between the two schooling regimes just described is charactersitic of developing economies where children can often act as productive units to complement the income receipts in the household. For example, when adverse stochastic shocks temporarely lowers the productive capacity of the adults in the household, the adjustment mechanism often consists on substituting the production of adults by that of children at the expense of shifting the children’s school regime to a non-school regime. Although the shift from a school to a non-school regime may be the result of some temporary adverse effect (e.g., a temporary illness of an adult in the household), the consequence in terms of human capital formation for the child – and therefore for the household – may become permanent. Likewise, the effect of some positive stochastic shock that enhances the household’s productive capacity could result in the decision to shift from a non-school to a school regime for the children in the household resulting in permanent gains in the form of human capital accumulation. In general, the effect of a stochastically changing environment implies that risk avers agents make investment decisions under uncertainty; in the case of developing economies a decision to send the children to school is an investement choice with the opportunity cost implied by the loss in production for the household. In the presence of uninsurable uncertainty the effect of contemporaneously optimal choices may lead to lower future gains in terms of human capital. This model developed in this paper provides a tool to measure endogenous human capital growth rates taking into account the uncertainty implied by a stochastically chaning environment. Identification and estimation of endogenous human capital formation at the household level requires dynamic data capable to identify endogenous productivitiy growth, i.e., per household income growth net from the effects that an stochastically changing environement has on the households capacity to produce and condioning on the dynamically changing characteristics and preferences in the household that ultimantely determine the household’s educational (investment) choices. The same model shows that human capital has two components. The first component refers to that part of human capital attributable to intergererational transfers, i.e., human capital past from parents to children and that is present even if the children in the household do not attend school. The second component refers to that part of human capital formation attributable to formal education and adds to human capital if children attend school. The quantification of endogenous human capital aims at providing evidence for the potential effect from policy interventions, i.e., the implementation of the model aims at quantifying the relative gains from investment on human capital (e.g., education) for children in households where the effect of uncertainty may lead to lower educational investment that may ultimately hinder their human capital formation and, consequently, lowers the economic growth for the country as a whole.

The implementation of the model is illustrated using the Kagera Health and Development study (1991-2004). This data provides demographic, social and economic information for the North-Western region of Tanzania. The nature of the data is particularly appropriate for the implementation of the model since it allows us to decompose the sample among different household types (i.e. from orphans to two parent households) as well as to distinguish among households according to productivity types. Furthermore, the data is dynamic and informative with regards to the schooling choices that parents make for their primary school age children: in Tanzania this implies children from the ages of 6 to 15. Estimates of human capital parameters suggest that orpahs are the ones to benefit most from educational choices, where orphans are defined as children living in households where both parents are dead and an older sibling has become the head of the household. We find that children in orphan households experience human capital growth rates that are 6.2% higher when these children are sent to school relative to orphans that are observed in a non-school regime. In the case of single households and two parent households, the effect of schooling on human capital is positive but remains small when compared to the effect that parents have on their children's human capital formation. For example, the estimates suggest that human capital formation increases by 81.5% over a period that covers the primary school age years for two parent households where the primary school age children are observed in a non-schooling regime. Sicen these children are observed as not attending school the implication is that human capital is attributable to intergenerational transfers only. The same estimate using households where their primary school age children are observed as attending school implies that in total (i.e., adding both the intergenerational transfers and the effect from formal education) the productive capacity increases by 80.4% over the same period. The difference between the two type of school regimes is small (1.1%) and is not statistically significant. Thus, it seems that parental presence is far more important then education at inducing human capital formation. In fact, we notice that the main difference regarding human capital formation is found between single and two parent households when the two regimes are compared purely on the basis of intergenerational transfers, i.e., when the comparison focuses on households in a non-schooling regimes so that any gain in terms of human capital can only be attributed to knowledge transfers from parents to children. In this case we estimate that human capital growth can be 33% higher for households where both parents are present relative to single parent households. This evidence seems to suggest that in the case of developing economies with a rural background (i.e., for household with productivity related to the agropecuarian industry), the presence of both parents is what makes a differnce to the children's potential productive capacity in periods ahead. Developing economies that are significally affected by epidemics (e.g., the effect of the AIDS epidemic) often experience a deterioration of their productive population as consequence of premature mortality. A consequence of this is a significant increase in the number of orphans. Our results seem to suggest two policy actions that may lead to a better outcome for children that become orphans as result of parental premature death: making sure these children attend school can help to make up for the human capital loss that comes from a lack of parental transfers in the form productive knowledge. Alternatively, initiatives that place orphan children in households where parents are present can also lower the loss implied by the lack of human capital transfers that results from a state of orphanhood. Having said these, we have to be cautions with the interpreatation of our results. The data so far employed in this illustration (i.e., The Kagera Health and Development study, 1991- 1994) is a high quality data set with the drawback that it provides informarion on a 6 month basis. Such short intervals of time imply that we cannot observe the evolution of families over very longer periods and this leads us to rely on the assumption that current cohorts are informative about households in periods ahead. Our estimates may pick up some of the effects of schooling but data sets that provide longer periods would be more informative on the effects of schooling on human capital formation by household type. Finally, it is fortunate to learn from the Kagera Health and Development Study (1991-1994) that in the North-Western part of Tanzania we cannot detect a large number of orphan households. Howerver, the low sample size for this type of households implies that sampling variance – or lack of it – does not provide a clear understanding on the significance of the estimates. Unfortunately, in other societies (e.g., in the Southern region of Sub-Saharan Africa) epidemics such as AIDS has lead to a more significant increase in the number of orphans. Assuming that the correct type of data exists for these societies, the same quantitative analysis as the one presented here may provide a more informative understanding on the effects that schooling has with regards to the human capital formation for children living in orphan households.

Reference

TO BE UPDATED

Angrist, J., and A. Kruger (1999), Empirical strategies in Labour Economics, in O. Ashenfelter and D. Card (eds.), *Handbook of Labour Economics*, Volume III, Amsterdam, 1277-1366

Ashenfelter, O., and D. Card (1985), Using the longitudinal structure of earnings to estimate the effect of training programs, *Review of Economics and Statistics*, 67, 648-660.

Auerbach, A. and L. Kotlikoff (1987), *Dynamic Fiscal Policy*, Cambridge University Press

Bell, B., R. Blundell, and J. van Reenen (1999), Getting the unemployed back to work: an evaluation of the New Deal proposals, *International Tax and Public Policy*, 6, 339-360.

Bell, C., S. Devarajan and H. Gersbach, (2003), The long run economic cost of AIDS: a model with an application to South Africa, *The World Bank Economic Review*, vol.20, no.1, pp 55-89

Bontemps, C., J.M. Robin, and G. van den Berg, (2000), An empirical equilibrium search model with search on the job and heterogeneous workers and firms, *International Economic Review*, 41 (2), 305-358

Browning, M., L. Hansen, and J. J. Heckman (1999), Micro data and General Equilibrium models', in Taylor and Woodford (eds.), *Handbook of Macroeconomics*, Vol. 1, Ch.8, Amsterdam: Elsevier Science, pp. 543-633

Costa-Dias, M., (2002), *The evaluation of social programmes*, PhD dissertation series, University College London.

Costa-Dias, M., R. Blundell, and C. Meguir, (2002), The impact on wage subsidies on education and employment: a general equilibrium approach, *IFS working paper*, mimeo.

Deaton, A. (1992), *Understanding consumption*, Oxford University Press.

Evans, D., and B. Jovanovic (1989), An estimated model of entrepreneurial choice under liquidity constrains, *Journal of Political Economy*, 97, 808-827

Froelich, M and R. Vazquez-Alvarez (2007), HIV/AIDS and HIV-knowledge: can information campaigns reduced the HIV infection? The case of Kenya, WP-SIAW-HSG, University of St.Gallen, Switzerland

Heckman, J.J., L. Lochner, and C. Taber (1998), Explaining rise inequality: explorations with a dynamic general equilibrium model of labour earnings with heterogeneous agents, *Review of Economic Dynamics*, 1, 1-58

Heckman, J.J, R. LaLonde, and J. Smith (1999), The economics and econometrics of Active Labour Market Programs, in Ashenfelter and Card (eds.), *Handbook of Labour Economics*, Vol. 3, Ch 31, Amsterdam: Elsevier Science, pp. 543-633

Heckman, J.J., L. Lochner, and R. Cossa (2003), Learning by doing versus on the job training: using variation induced by the EITC to distinguish between models of skill formation, in E. Phelps (eds.), *Designing inclusion: Tools to raise low-end pay and employment in private enterprise*, Cambridge University Press, pp. 74-130

Heckman, J.J., and J. Smith (1999), The pre-program earnings dip and the determinants of participation in a social program: implications for simple program evaluation strategies, *The Economic Journal*, 109, 313-48.

Huggett, M. (1997), The one-sector growth model with idiosyncratic shocks: steady-state and dynamics, *Journal of Monetary Economics*, 39, 385-403.

Huggett, M., and G. Ventura (1999), On the distributional effects of social security reform, *Review of Economic Dynamics*, 2, 498-531.

Kihlstrom, R., and J.J. Laffont (1979), A general equilibrium entrepreneurial theory of firm formation based on risk aversion, *Journal of Political Economy*, 87, 719-990.

Lechner, M and R. Vazquez-Alvarez (2007), Stochastic labour market shocks, labour market programmes and human capital formation, WP SIAW-HSG, University of St.Gallen, Switzerland

Lopez, A., C. Mathers, M. Ezzati, D. Jamison and C. Murray (2006), Burden of disease and risk factors, Oxford University Press and the World Bank, Washington

Machin, S., and A. Manning, (1999), The causes and consequences of long term unemployment in Europe, in Ashenfelter and Card (eds.), *Handbook of Labour Economics*, Vol. 3, Ch.47, Amsterdam: Elsevier Science, pp. 3085-3139.

Magnac, T., and J. M. Robin (1991), A dynamic model of choice between wage work and self-employment with liquidity constrains, CREST Working Paper, No. 9116, Paris, France.

Magnac, T., and J. M. Robin (1996), Occupational choice and liquidity constrains, *Ricerche Economiche*, 50, 105-133.

Nunziata, L. (2003), Labour market institutions and the cyclical dynamics of employment, *Labour Economics*, 10, 31-53

Perroni, C., (1995), Assessing the dynamics efficiency gains of tax reforms when human capital is endogenous, *International Economic Review*, 36, 957-25.

Regerson, R., R. Shimer, and R. Wright, (1995), 'Search theoretic models of the labour market: a survey', *Journal of Economic Literature*, 43, 959-988.

Stokey, N. and L. Robert (1989), Recursive methods in Economic Dynamics, CUP