

On the Origin and Causes of Economic Growth

Nicolas Roys and Ananth Seshadri*

Department of Economics, University of Wisconsin-Madison

March 2014

Abstract

This paper presents a model of human capital accumulation to better understand the take-off from stagnation to growth from 1500 to 2000 AD. Finitely lived households choose the quantity and quality of children. The key ingredient of the model is a spillover from parents to children in the accumulation of human capital and an economywide spillover. Depending on the size of the spillovers, the model can generate protracted transitions. Starting from an initial level of human capital, the economy can take *centuries* to reach 95% of the steady state output per capita with a half-life of around 250 years. The model can rationalize the demographic transition as well as the industrial revolution without resorting to exogenous changes in productivity. It is consistent with the changing cross-sectional relationship between income and fertility as well as the decline in the concentration of wealth. Macro evidence on convergence lends support to our formulation.

*We thank seminar participants at the NBER Macroeconomics Across Time and Space Workshop, Chicago Fed, Wisconsin and the European Economic Association Meetings at Gothenburg (Sweden) as well as Robert Barro, Robert Barsky, Gadi Barlevy, Hal Cole, Dean Corbae, and Jeremy Greenwood for very helpful comments.

1 Introduction

The process of economic development is featured by a very long period of stagnation followed by rapid take-off (see Figure 1 for GDP per capita in a select few countries from 1500 to 2000 AD). The three industrial revolutions, the first in the United Kingdom and the second and the third in the United States have had a profound effect on the living standards of a large fraction of the world's population. There is a very large literature that seeks to understand the process through which growth occurs. Understanding this phenomenon has large consequences for welfare. Broadly speaking, there are two strands of literature that are inextricably intertwined. One line of work seeks to identify the 'proximate causes' of economic growth. Various mechanisms have been proposed and the most influential ones that fall in this category include physical capital (Solow, 1956, Cass, 1965, Koopmans, 1965), human capital (Lucas, 1988, Becker, Murphy and Tamura, 1988 on the theoretical side and Mankiw, Romer and Weil, 1992 and Barro, 1997 on the empirical side), and innovation (Romer, 1990; Grossman and Helpman, 1991; and Aghion and Howitt, 1992). In the spirit of the standard Neoclassical production function, these mechanisms help us further our understanding of the effects of exogenous shifts on productivity, physical capital and human capital.

Another line of work seeks to identify the 'fundamental origins' of growth. The fundamental origins include institutions (North and Thomas, 1973, Acemoglu, Johnson and Robinson, 2001) and geography (Sachs, 2001). These papers argue that factors that represent proximate causes (such as R&D, physical capital, human capital) "are not causes of growth; they *are* growth" (North and Thomas, 1973). That is, they argue that these fundamental factors described above serve as driving forces which influence the growth processes through the mechanisms identified as proximate causes. Whether institutions cause growth (human capital) or more educated populations lead to the creation of better institutions is the subject of considerable recent debate. Recent work by Glaeser et. al. (2004) and Barro (2012) argue that human capital is a fundamental cause of economic development - that is, they view better institutions as the result of greater human capital acquisition.

Many believe that human capital is central to the growth process - either as an origin

or as a cause. Over the long run, human capital accumulation gained prominence and led societies from stagnation to a period of remarkable growth. We begin by delving deeper into models of human capital. Perhaps the most appropriate place to begin on any discussion of human capital and growth is the seminal article by Lucas (1988) in which he emphasizes aggregate externalities to human capital accumulation and also assumes that the technology to produce human capital is linear. While influential judging by the enormous body of work on endogenous growth, this work has not been subject to much quantitative scrutiny. There are two concerns raised against this line of investigation. First, what is the empirical justification for the linearity (AK) assumption? Solow (1994) in his review of the literature argues that “The conclusion has to be that [the constant returns] version of the endogenous-growth model is very un-robust. It cannot survive without exactly constant returns to capital. But you would have to believe in the tooth fairy to expect that kind of luck.” Second, how does one accurately measure the aggregate spillover. The aggregate spillover is not an easy object to estimate and many existing studies do not find much evidence for it. Estimates of marginal social returns to investing in human capital are typically not very different from estimates of marginal private returns.

In other related work that emphasizes human capital and economic development, Becker, Murphy and Tamura (1990) present a model featuring multiple steady states and assume that the return to accumulating human capital increases with the stock of human capital. Exogenous shocks (luck) move the economy from stagnation to a period of growth. More recently, Galor (2005) argues that a successful theory of development must provide a unified explanation for the entire time series path over the very long run. The large body of work has primarily been theoretical and has not been the subject of much quantitative scrutiny. The natural next step is to understand long run growth through the eyes of a model consistent with micro-level evidence on human capital accumulation (without resorting to ad-hoc assumptions of linearity or arbitrary assumptions on the significance of the aggregate externality) and to examine the quantitative implications.

In this paper, we do just this. We start with a fairly standard life-cycle model of human capital accumulation a la Ben Porath (1967). This model has been widely studied in the literature and estimates of this model are readily available. The life-cycle model generates quick transition to the steady state, a property well known to students of growth theory.

We embed this in a dynastic setting and most importantly, we add an intergenerational feature. We allow for a spillover of human capital from the parent to the child. This considerably alters the convergence properties of the life-cycle model and sets the stage for long protracted transitions. We embed this in a heterogeneous agent model with endogenous fertility along the lines of the influential work of Barro and Becker (1989) in which altruistic parents internalize this spillover. The industrial revolution and demographic transition have long been thought to have been inextricably intertwined and rationalizing the timing and magnitudes of these changes is an important endeavor. Over time, starting from an initial level of human capital, fertility rates decline. The decline in birth rates leads to lower real interest rates, since physical capital rises by more than human capital. The primary effect at play is an increase in bequests by the wealthiest households. Lower interest rates eventually spur human capital accumulation and lead to very rapid growth after parents become sufficiently wealthy. Initially, only the rich invest efficiently in the human capital of their children but as time goes by, more and more parents invest in the human capital of their children.

We ask whether this parsimonious framework can account for the observed pattern of growth over the last 500 years when we start the model at an initial condition in 1500 AD and assume no exogenous change whatsoever. Quite remarkably, we find that it can. The model generates a very long period of stagnation depending on the size of the parental spillover. The model implies slow convergence to the steady state; convergence from an arbitrary initial level of human capital can take as many as centuries with a half-life of around 250 years. The model thus generates dynamics that are capable of rationalizing long run development paths. Initial conditions can have a rather large influence on when development occurs. Interestingly, the model's cross-sectional predictions also match up with the evidence. In circa 1700, the model implies that rich parents have more surviving children than the poor while in the modern era, this correlation reverses. This is precisely what we observe in the data. Furthermore, the model's predictions for the concentration of wealth over time, which declined considerably, is also consistent with available evidence.

The parental spillover term is critical in generating the observed pattern. Being able to defend this parameter, empirically, is important. To do this, we look at both macro time series evidence as well as micro evidence. There has been considerable interest in deter-

mining whether or not economies converge and if so, to what extent. Barro (2012) reviews the evidence and concludes that the iron law of convergence holds for the past 100 years - countries eliminate gaps of real GDP per capita at the rate of 2% per year. We use time series data for the past 120 years and demonstrate that the calibrated spillover parameter implies a convergence rate very similar to the iron law. Furthermore, we also show that smaller values of the spillover parameter imply much faster rates of convergence. Finally, we also use micro evidence to estimate the parental spillover term. While calculating the social return to education is no easy task (something that would be required to say defend the aggregate spillover parameter), the parental spillover term can easily be recovered by examining estimates of the effect of parental schooling on child earnings, controlling for the child's schooling level. To be clear, we look at individuals who acquire the same years of schooling but have parents with different human capital levels (schooling). The evidence suggests that parents have a very large effect on the wages of children with the *same* schooling level. We use our theory to estimate the size of the parental spillover term which corresponds rather closely to the value obtain from a macro level calibration. We conclude that the parameter values that are instrumental in generating the long period of stagnation followed by rapid growth are empirically defensible.

The rest of the paper is organized as follows. Section 2 lays out the key model ingredients. It starts off with a standard life-cycle model of human capital accumulation and adds to it a parental spillover. It casts this in a model of endogenous fertility and endogenous lifespans. Section 3 describes model calibration. Section 4 reports the main results. The model generates a long transition starting from some initial human capital level. Since the model features heterogeneity, this section also presents the cross-section implications for fertility choice and wealth holdings as they evolve over the past few centuries. Section 5 examines the implications for the iron law of convergence. It also examines the implications for alternative models for convergence. Section 6 presents micro evidence on the impact of parental education on the earnings of children to estimate the parental spillover term. Section 7 concludes.

2 The Model Economy

2.1 Model Overview

In this section, we lay out the key model elements. There are a few different pieces to the model and we describe each of these in detail. The objective is to construct an economic framework that can generate long transitions by modeling the period of stagnation for centuries and the subsequent take-off in the 1800s while simultaneously being consistent with time series evidence on schooling, fertility rates and lifespan. The model we present below is an amalgam of a life-cycle model and a dynastic framework. Consequently, we need to specify preferences over the life-cycle as well as over generations. We assume that individuals care about their utility while alive as well as the utility of their progeny. Standard life-cycle models with physical capital and human capital feature rapid convergence to a steady state (see Barro and Sala-i-Martin, 2004). Even with a very high capital share, these economies possess half-lives that are much smaller than what would be required to generate the slow transition from stagnation to growth. We augment an otherwise standard life-cycle human capital accumulation equation with a spillover from parent to child - we assume that the human capital of the parent is a productive input in the production of human capital for the child. The parent takes into account the impact on future generations' human capital accumulation through his altruism. Along the transition from some initial condition to the eventual steady state, real interest rates decline. This decline eventually drives economic development.

We take seriously the view that a successful theory must be able to generate the fall in real rates of return. There are at least two ways to model changes in real rates of return. Models with financial intermediation can generate this pattern. An alternative is to consider an endogenous fertility framework along the lines of the pioneering work of Barro and Becker (1989) - these models have the ability to generate declines in real interest rates when fertility falls. We proceed with the latter framework; the industrial revolution and the demographic transition have long been thought to be closely related. Generating the fall in interest rates as well as the decline in fertility and the ensuing rise in GDP and schooling starting from an initial condition without resorting to any exogenous change is the overriding goal of the model presented here. Clearly, both physical and human capital

have increased in importance over the last few centuries.

Preferences: Since the objective is to move away from standard life-cycle decision problem and consider intergenerational flows from parents to children, we draw on the dynastic model presented by Barro and Becker (1989). Parents are altruistic towards the well-being of their dynasties. Parents can invest in the human capital of their children and also leave bequests.

Fertility and Discounting: As in Barro and Becker (1989), we assume that the number of children, f_t affects the effective discount factor of the parent, $b(f_t)$. Specifically, the weight that the parent places on his future generations is an increasing and concave function of the choice of the number of children. As is well known, in a steady state, there is a one to one relationship between the real interest rate and the fertility rate.

Physical Capital: The stock of physical capital is the aggregate of life-cycle savings and intergenerational transfers. Modeling both of these is important. Before 1850, when lifespans were short, bequests were the significant portion of capital accumulation. Indeed, available evidence suggests that wealth was very concentrated and bequests were prevalent amongst the very wealthy. During this period of time, human capital accumulation was not given much importance and since individuals worked until their death, life-cycle savings were also not very relevant. As economies got richer, households engaged in more savings both in the form of bequests and more importantly, saving for retirement as life expectancies rose.

Human Capital: Since this is the central feature that helps generate slow transition to the steady state we will discuss this aspect in greater detail. There are numerous models of human capital accumulation that have been studied in various fields of economics. Some of the most prominent of these are Arrow (1962), who analyzes a technology featuring learning by doing, Becker and Tomes (1976), who emphasize intergenerational transmission of human capital, Ben Porath (1967), who lays out a life-cycle model of human capital accumulation (Mincer, 1971 is a special case of Ben Porath, 1967 where the post schooling investment is assumed to be linear) and Lucas (1988), who specifies an infinite horizon model of human capital accumulation in which the technology is assumed to be linear and features external effects. We start by describing the Ben-Porath (1967) model before turning to how we augment it with a parental spillover. This is the workhorse of modern

labor economics and has been used widely in various contexts - the effects of rising skill premia, cross country income differences, the study of inequality within a country and many other applications. The individual begins life (at age 6) with his initial stock of human capital, h_0 and his ability level, z_h . Human capital is produced using time n , the pre-existing stock of human capital h along with goods inputs x . The technology with which human capital is produced is given by $z_h[nh]^{\gamma_1}(x)^{\gamma_2}$ where each individual is assumed to be born with the same human capital level, h_0 . It is well known that such an economy features a rather rapid transition to a steady state.

In this framework, the time path of n decreases with age. Assuming decreasing returns to scale, i.e. $\gamma_1 + \gamma_2 < 1$, if the initial stock of human capital is low enough, the time allocation decision is constrained at 1 for a few years and then declines. The period of time during which n equals 1 is labeled the schooling period. Finally, the intermediate input x stands in for all educational inputs aside from student time. These inputs are also produced using physical capital and human capital.

This technology is usually cast in a decision problem that is finite horizon problem. When the individual retires, his stock of human capital depreciates completely. There is no transmission of human capital from one generation to the next. This framework cannot generate the rise in schooling or GDP without resorting to exogenous changes in productivity. Furthermore, given that each subsequent cohort begins life with the same initial stock of human capital, convergence to the steady state from some initial condition say due to changing life expectancy or real rates of return is rather quick. This leads us to think further about inter-generational linkages. Clearly, parents have an influence on both the amount of learning that happens before school entry (h_0) as well as the learning ability of the child (z_h).

We will now augment the model with a spillover from parents to children. We do this in the simplest possible manner - we assume that the human capital of the parent affects initial human capital and augments ability. The impact of parental human capital on children's earnings potential has been recognized at least as early as Marshall (1890). In an insightful comparison of the children of unskilled laborers with the children of an artisan, Marshall lays out the significant advantages conferred on those who are born into the 'higher grades of society'. The child of the artisan is brought up in more refined home

and his parents are likely to possess more human capital. He goes on to add (Vol. 1, page 592) that “the most valuable of all capital is that invested in human beings; and of that capital the most precious part is the result of the care and influence of the mother”.

Parents have a large influence on the knowledge of a child at age 6. They also influence their ability to learn throughout their lives. We model these effects in a parsimonious manner by augmenting the standard human capital technology presented above with parental human capital, h^p as a productive input. We assume that the human capital production is given by

$$z_h (nh)^{\gamma_1} x^{\gamma_2} (h^p)^{\gamma_3}$$

and the human capital of the child at age 6 is given by

$$(h^p)^{\gamma_4}.$$

The intergenerational spillover terms are governed by γ_3 and γ_4 . When the child is born, the parent’s human capital at the time of birth spills over to his child. The parent knows this and invests in his own human capital accordingly by internalizing the spillover effect. A higher human capital parent is better able to transmit human capital to his child and is *also* better able to augment learning during the formative first 6 years of life. This is a rather standard assumption in the literature on intergenerational transmission. See, for instance, Becker and Tomes (1976). Notice that the problem presented here can be viewed through the lens of the standard human capital framework with two normalizations: call effective ability $z_h (h^p)^{\gamma_3}$ and relabel the initial stock of human capital $(h^p)^{\gamma_4}$. While standard human capital theory assumes that ability and initial stock of human capital are parameters, we assume that these are influenced by parental human capital.

Heterogeneity: Modeling this complex transformation requires us to deviate from the homogeneous agent assumption. A single agent model will imply that bequests will be zero for a while and then as the economy gets richer, bequests become positive. Such a prediction will be inconsistent with the evidence that bequests were the most significant portion of nonhuman wealth centuries ago and then gradually declined in significance. Being born to a wealthy family is much less important now than it was in say 1500 AD. We introduce heterogeneity in its simplest form: we assume that dynasties differ permanently in terms of innate ability z_h . We also assume there is no uncertainty. Learning ability is

transmitted perfectly from parent to child and there is no socioeconomic mobility. The paper is primarily concerned with the computation of transition paths over long periods of time, a computationally demanding problem. We abstract from adding noise to the ability transmission process. Incorporating heterogeneity allows us to examine the predictions of the model along various dimensions such as the cross-sectional relationship between fertility and income as well as the concentration of wealth and how these relationships changed over time. More important, it allows for dynastic linkages (positive bequests) to occur slowly and at different points in time for different types and this helps rationalize the slow transformation from stagnation.

Life Span: Clearly, lifespans are a function of economic development. We allow for lifespans to vary with the level of development. We assume that households can invest l units of goods in their children when they are born and these l units determine the lifespan of their children through the function $T(l)$. To keep things manageable, we do not model investment in health capital over the life cycle as that would add another state variable.

2.2 Environment

Before proceeding to describe the environment, a word on notation is in order. Throughout the paper, t stands for calendar year of birth, i stands for age, and j stands for ability type. For any variable x , x_t^{ij} denotes the relevant variable for an individual born in year t with ability type j who is currently age i , and x_t^{ijk} denotes the relevant variable x for the child (k for kid) of an individual born in year t with ability type j who is currently age i (assuming an individual becomes a parent at age B , the age of the child is thus $i - B$). When a variable x does not depend on age (such as fertility, f) we refer to x_t^{ij} as x_t^j . When there is no confusion, we also use x_t^{ij} and $x_t^{i,j}$ interchangeably. We do the same for x_t^{ijk} .

The economy is populated by overlapping generations of individuals who live for $T(l_t^j)$ periods. The (discrete) timeline proceeds as follows. After birth at time t , an individual remains attached to his parent until he is I years old; at that point he creates his own family and has, at age B , f_t^j children that, at age $B + I$, leave the parent's home to become independent. He works until retirement, R_t^j . The variables I and B are exogenous and time-invariant. R_t^j is allowed to vary over time according to the variation in lifespans.

The only source of heterogeneity across individuals is z_h^j . where j indexes an ability type. As stated above, we assume that this learning ability is perfectly transmitted across generations. In our computational experiment, we will assume there are 5 different values for z_h^j with $j = 1, \dots, 5$.¹

The dynamic programming problem of a young adult born at time t with ability j who possesses h_t^{Ij} units of human capital at the time he becomes independent I , and initial wealth (a bequest from his parents which is most appropriately thought of as a lifetime intergenerational transfer net of any support that the child provides to the parent) equal to b_t^{Ij} , at age I , is given by choice of variables consumption over the life-cycle $\{c_t^{ij}\}_{i=I}^{T_t^j}$, investments in human capital on the job $\{x_t^{ij}, n_t^{ij}\}_{i=I}^{R_t^j}$, consumption of their children $\{c_t^{ijk}\}_{i=B}^{B+I-1}$, schooling time and expenditures $\{n_t^{ijk}, x_t^{ijk}\}_{i=B+6}^{B+I-1}$, fertility rate f_t^j , bequests $b_t^{B+I,jk}$ and expenditures on life extension for their children l_t^{Bjk} to solve

$$V_t(h_t^{Ij}, b_t^{Ij}, h_t^{jp}, z_h^j) = \max \left[\begin{array}{c} \sum_{i=I}^{T_t^j} \beta^{i-I} u(c_t^{ij}) \\ + b(f_t^j) \left(\sum_{i=B}^{B+I-1} \beta^{i-I} u(c_t^{ijk}) + \beta^B V_{t+B}(h_t^{B+I,jk}, b_t^{B+I,jk}, h_t^{Bj}, z_h^j) \right) \end{array} \right]$$

¹The model economy presented here is related to Manuelli and Seshadri (2009) who present a model of fertility choice and human capital accumulation. There are at least three critical differences: first, the model presented here features a parental spillover which considerably alters the convergence properties of the model. Second, we add heterogeneity in ability. Third and most importantly, we compute transition paths here while the focus in that paper was on steady states.

subject to a lifetime budget constraint

$$\begin{aligned}
& \sum_{i=I}^{T_t^j} \frac{c_t^{ij}}{\prod_{m=I}^i (1+r_{t+m})} + \sum_{i=I}^{R_t^j} \frac{x_t^{ij}}{\prod_{m=I}^i (1+r_{t+m})} + \\
& f_t^j \left[\overbrace{\sum_{i=B}^{B+I-1} \frac{c_t^{ijk}}{\prod_{m=I}^i (1+r_{t+m})} + \sum_{i=B+6}^{B+I-1} \frac{x_t^{ijk}}{\prod_{m=I}^i (1+r_{t+m})} + \frac{b_t^{B+I,jk}}{\prod_{m=I}^{B+I} (1+r_{t+m})} + \frac{l_t^{Bjk}}{\prod_{m=I}^B (1+r_{t+m})}}^{\text{Cost of a Child}} \right] \\
& = \underbrace{\sum_{i=I}^{R_t^j} w_{t+i} \frac{h_t^{ij}(1-n_t^{ij})}{\prod_{m=I}^i (1+r_{t+m})}}_{\text{Lifetime Earnings}} + \underbrace{f_t^j \sum_{i=B+6}^{B+I-1} w_{t+i} \frac{h_t^{ijk}(1-n_t^{ijk})}{\prod_{m=I}^i (1+r_{t+m})}}_{\text{Child Labor Earnings}} + \underbrace{b_t^{Ij}}_{\text{Bequest}},
\end{aligned}$$

the equation governing the lifespan of their children

$$T_t^{Bjk} = T(l_t^{Bjk}), \quad (1)$$

the evolution of human capital for the parent

$$h_t^{i+1,j} = z_h^j [n_t^{ij} h_t^{ij}]^{\gamma_1} (x_t^{ij})^{\gamma_2} (h_t^{jp})^{\gamma_3} + (1 - \delta_h) h_t^{ij}, \quad i \in [I, \dots, R_t^j], \quad (2)$$

where

$$h_t^{jp} = h_{t-B}^{Bj} \quad (3)$$

the evolution of human capital for the child

$$h_t^{i+1,jk} = z_h^j [n_t^{ijk} h_t^{ijk}]^{\gamma_1} (x_t^{ijk})^{\gamma_2} (h_t^{Bj})^{\gamma_3} + (1 - \delta_h) h_t^{ijk}, \quad i \in [B+6, \dots, B+I-1], \quad (4)$$

a constraint on the time allocation decisions,

$$0 \leq n_t^{ij}, n_t^{ijk} \leq 1, \text{ for all } t, i, j \quad (5)$$

the child's initial stock is given by

$$h_t^{B+6,jk} = (h_t^{Bj})^{\gamma_4} \quad (6)$$

and a non-negative restriction on bequests

$$b_t^{B+I,jk} \geq 0.$$

To solve the model, an initial condition needs to be specified. Let the initial distribution across agents be denoted by $\Psi_j(h_0^j, a_0^j, h_0^{jp}; z_h^j)$ where a_0^j stands for assets

The value functions are indexed by time since time-varying interest rates affect individual decisions. The first two terms in the lifetime budget constraint are expenses related to own consumption and investment in job training for the adult. The term inside the square bracket represents the cost per child. We assume that there are no life-cycle borrowing constraints and this allows us to write down a life-cycle budget constraint. We however assume that parents are unable to borrow against the future incomes of their children, that is bequests are assumed to be non-negative. This restriction is the only market friction in our model. As will be clear from the ensuing analysis, it plays a very important role. It leads human capital decisions to be inefficient for a majority of the households. It is not easy even in contemporary society for parents to use the ability of their children as collateral to obtain a loan and have their children legally responsible for debts incurred on their behalf. The restriction on bequests captures, in a reduced form way, such inefficiencies.

Parents are imperfectly altruistic. They value the utility of their children but place a different weight on the utility of their children relative to their own utility. In our simulations, $b(f) = \alpha_0 f^{\alpha_1}$ captures the degree of altruism. If $\alpha_0 = 1$, and $\alpha_1 = 0$, the preference structure mimics that of the standard infinitively-lived agent model. Values of α_0 less than 1 capture the degree of imperfect altruism. The term $V_{t+B}(h_t^{B+I,jk}, b_t^{B+I,jk}, h_t^{Bj}, z_h^j)$ stands for the utility of the child at the time he becomes independent.

What are the conditions under which endogenous growth obtain? Along a balanced growth path, time allocation decisions such as schooling remain constant while stocks of human capital grow at a constant rate. Endogenous growth obtains when the human capital production functions are linear, that is $\gamma_4 = 1$ and $\gamma_1 + \gamma_2 + \gamma_3 = 1$. These assumptions are fairly standard in the endogenous growth literature but empirical significance of it has been the subject of considerable debate. It is interesting to note that Lucas (1988) defends his choice of a linear production function (equation 13 in his analysis) in an infinite horizon context based on a model of life-cycle wage growth.

"It is a digression I will not pursue, but it would take some work to go from a human capital technology of the form (13), applied to each finite-lived individual (as in Rosen's theory), to this same technology applied to an entire infinitely-lived typical household or family. For example, if each individual acquired human capital as in Rosen's model but if none of this capital were passed on to younger generations, the household's stock would (with a fixed demography) stay constant. To obtain (13) for a family, one needs to assume both that each individual's capital follows this equation and that the initial level each new member begins with is proportional to (not equal to!) the level already attained by older members of the family."

Note that Lucas makes specific assumptions on γ_3 and γ_4 . We depart from Lucas's suggestion in three significant respects. First, rather than assuming a linear technology for the life-cycle of a household's decision problem, we will let evidence on wage growth over the life-cycle and schooling pin down the returns to scale on the human capital production function. Second, Lucas's suggestion of having the initial stock of human capital being proportional to parent's human capital also assumes linearity. We posit a dependence of the child's initial human capital on the parent's human capital level and allow the data speak to whether 'proximity to linearity' helps the model match up to time series evidence as well as cross-sectional implications. Third, Lucas emphasizes externalities while our model does not feature aggregate external effects. Our focus is on the impact of parental human capital on the child's human capital investment decisions.

If bequests are in the interior, then the FOC for bequests at a steady state reads

$$u_1(c^{Ij}) \frac{f^j}{(1+r)^B} = b(f^j) \beta^B V_2(h^{B+I,jk}, b^{B+I,jk}, h^{B+I}; z_h^j),$$

while the envelope condition associated with the parent starting off with a slightly higher human capital is

$$V_2(h^{Ij}, b^{Ij}, h^{jP}) = u_1(c^{Ij}).$$

Combining the two conditions, we get

$$u_1(c^{Ij}) \frac{f^j}{(1+r)^B} = b(f^j) \beta^B u_1(c^{B+I,jk}).$$

Assuming that the altruism function and the utility functions are both of the power function form ($u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ and $b(f) = \alpha_0 f^{\alpha_1}$), we get

$$\frac{(f^j)^{1-\alpha_1}}{\alpha_0} = [\beta(1+r)]^B \left(\frac{c^{B+I,jk}}{c^{Ij}} \right)^\sigma.$$

At a steady state, the value functions of the parent and the child coincide and consumption remains constant across generations. Hence, we arrive at the familiar condition that the interest rate equals the rate of time preference and fertility does not vary by ability type j . As economies grow along the transition to the steady state, fertility rates and real rates of return fall. Indeed both forces (fertility rates and growth rates) work in the same direction leading to declining interest rates and higher human capital accumulation.

If bequests are positive, investments in human capital are efficient and in the absence of the parental spillover terms, they coincide with the optimal stocks from standard human capital theory. At the steady state, real interest rates and fertility rates are related one for one. Fertility rates effectively appear in the discount factor. If the economy is poor enough, the parent may well want to borrow against their childrens' income. This leads bequests to equal zero and human capital investment decisions to be inefficient. We proceed with the following strategy. We start the economy off at an initial distribution across agents and analyze the transition to a steady state. To be a successful theory of economic development, the model must generate a fall in fertility starting around 1800. It should feature relatively stagnant GDP growth until then and the take-off would begin around 1800 when it would feature a rise in schooling.

2.3 Equilibrium

To determine aggregates, we need to determine the age structure of the population. We now describe the evolution of demographic structure and then describe how prices are determined.

Demographics It is useful to take stock of how demographics evolve. Recall that t denotes the year of birth while j stands for type. At any given point in time, individuals from different cohorts co-exist and the evolution of the population distribution is governed

by the fertility rate of cohorts as well as the life span of cohorts. Let μ_t^j denote the mass of individuals of ability level z_h^j born at time t . Let f_t^j denote the fertility rate of individuals of ability level z_h^j born at time t . Below is a table that depicts the cohorts that are alive at a given point in time.

Cohort	Fertility	Size
t	f_t^j	μ_t^j
$t + 1$	f_{t+1}^j	μ_{t+1}^j
\vdots		
$t + B$	f_{t+B}^j	$\mu_{t+B}^j = f_t \mu_t^j$
$t + B + 1$	f_{t+B+1}^j	$\mu_{t+B+1}^j = f_{t+1} \mu_{t+1}^j$

It is immediate that the condition

$$\mu_{t+B}^j = f_t^j \mu_t^j$$

holds. At the steady state the fertility rate f_t^j is constant across time t and ability type j and so $f_t^j = f$. Life span T^j is constant over time but varies with type j . The mass of individuals of ability type j at the steady state is given by

$$\mu_t^j = c_0^j (f)^{c_1 t}$$

where c_0^j and c_1 are constants to be determined. Substituting the above in the equation for evolution of the population, we get that $c_1 = \frac{1}{B}$. Hence, the mass of individuals of ability level z_h^j born at time t is given by $\mu_t^j = c_0^j (f)^{\frac{t}{B}}$. Let N_t^j denote total population of ability type j at time t . Then

$$N_t^j = \sum_{i=0}^{T^j} \mu_{t-i}^j = c_0^j (f)^{\frac{t}{B}} (f)^{-\frac{T^j}{B}} \frac{(f)^{\frac{T^j+1}{B}} - 1}{(f)^{\frac{1}{B}} - 1}.$$

Let N_t denote total population at time t . Since f is constant across types,

$$N_t = (f)^{\frac{t}{B}} \sum_{j=1}^5 c_0^j (f)^{-\frac{T^j}{B}} \frac{(f)^{\frac{T^j+1}{B}} - 1}{(f)^{\frac{1}{B}} - 1}.$$

Finally, the share of ability type j at the steady state, n^j is given by

$$n^j = \frac{N_t^j}{N_t} = \frac{c_0^j (f)^{-\frac{T^j}{B}} \left((f)^{\frac{T^j+1}{B}} - 1 \right)}{\sum_{l=1}^5 c_0^l (f)^{-\frac{T^l}{B}} \left((f)^{\frac{T^l+1}{B}} - 1 \right)}.$$

Aggregation To compute output it suffices to calculate the aggregate amount of human capital effectively supplied to the market, and the physical capital-human capital ratio. It is useful to define the oldest cohort alive at a particular point in time. Let o_t^j denote the oldest cohort alive of type j at time t .

$$o_t^j = \max\{i \in \mathbb{N} : i \leq T_{t-i}^j\}.$$
²

In every period t , the total stock of human capital allocated to market production, H_t , is calculated as the sum of the stocks of human capital supplied to the market by individuals who were born between time $t - o_t^j$ and $t - 6$ and is given by

$$H_t = \sum_{j=1}^5 \sum_{i=I}^{o_t^j} h_{t-i}^{ij} (1 - n_{t-i}^{ij}) \mu_{t-i}^j + \sum_{j=1}^5 \sum_{i=B+6}^{B+I} h_{t-i}^{ijk} (1 - n_{t-i}^{ijk}) \mu_{t-i}^j f_{t-i}^j.$$

The aggregate stock of physical capital, K_t equals life-cycle savings and intergenerational transfers³ across all ability types and is given by

$$K_t = \sum_{j=1}^5 \sum_{i=I}^{o_t^j} a_{t-i}^{i+1,j} \mu_{t-i}^j.$$

These stocks serve as an input into the final goods production function.

Equilibrium A Neoclassical firm produces output using a constant returns to scale production function $F(K_t, H_t)$. It rents physical capital, K_t and human capital, H_t from consumers. Optimization on the part of the firm implies that

$$(r_t + \delta_k) = F_k(\kappa_t, 1), \tag{7}$$

where $\kappa_t = K_t/H_t$ is the physical capital - human capital ratio and δ_k is the rate of depreciation of physical capital. The wage rate per unit of human capital, w_t , is,

$$w_t = F_h(\kappa_t, 1). \tag{8}$$

Then, aggregate output

$$Y_t = F(\kappa_t, 1)H_t.$$

We emphasize that there is no exogenous productivity change.

³ $a_{t-i}^{i+1,j}$ is the optimal choice of assets for the following year chosen by an individual born at time $t - i$ who is i years old at time t and of type j . These assets come from the year by year budget constraint of the individual.

Closing the Model Finally, given that the economy is closed, $C_t + I_t + X_t = Y_t$, where C_t stands for aggregate consumption, I_t is aggregate investment in physical capital, and X_t is aggregate investment in human capital. Physical capital evolves according to $K_{t+1} = (1 - \delta_k)K_t + I_t$.

3 Calibration

We use fairly standard functional forms. The production function for consumption goods is assumed to be Cobb-Douglas and takes the form

$$F(K, H) = K^\theta H^{1-\theta}.$$

We set θ to 0.33, a commonly used value. We assume that the period utility function is of the constant relative risk aversion variety and is given by

$$U(c) = \frac{c^{1-\sigma}}{1-\sigma}.$$

We set $\sigma = 0.5$ and the discount factor β to 0.96. Life span is given by the function $T(l) = \bar{T}(1 - e^{-\mu l})$ where the maximum possible lifespan \bar{T} is set to 90. The depreciation rate on physical capital is set at $\delta_k = 0.075$. A standard value for the capital output ratio in the United States is 2.52. This yields an interest rate of 5.5%. Available evidence suggests that in the United States, wage rates do not fall at the end of the life-cycle. This implies that the depreciation rate $\delta_h = 0$. In addition, we set $I = 18$, $B = 25$ and $R_{2000} = 64$. Along the transition, we set R_t to be the smaller of 64 or the endogenously chosen lifespan. We are left with the human capital technology parameters $\gamma_1, \gamma_2, \gamma_3, \gamma_4$, the health technology parameter, μ , the parameters governing the degree of altruism α_0, α_1 , the heterogeneity in innate ability z_h and finally the initial distribution across agents.

Our strategy is to choose the initial distribution as well as the other parameters so that the model mimics moments in year 2000. To compute the transition path, we will need to specify an initial distribution across agents over all these variables in addition to an initial population distribution. One possibility is to try and approximate this from data - unfortunately such information is not available. We take an alternate approach by

solving a life-cycle model ($\alpha_0 = 0$) with an exogenously specified interest rate of 10%, an exogenous ($\mu = 0$) lifespan of 40 ($\bar{T} = 40$), a fertility rate (f) for the five different types of 5, 5.5, 6, 6.5 and 7 and an initial human capital level (h_0) that is chosen so that the ratio of output per capita in 1500 to that in 2000 corresponds to the same object in the data. We use the resulting distribution obtained from solving this life-cycle economy as well as the stationary population distribution of this life-cycle economy as an initial condition. Note that if bequests are zero (which they are in our model economy in 1500), the allocations can be reasonably approximated by a corresponding life-cycle economy.⁴ Given the initial distribution across agents, we then proceed to solve for the transition path for a given set of parameter values. Nothing exogenous varies over time, in particular there is no change in productivity over time. Our calibration strategy involves choosing the parameters of our model so that the implications of the model economy presented above are consistent with observations for the United States for circa 2000 which is a point along the transition to the steady state.

The parameters z_h , γ_1 , and γ_2 are pinned down using observations on schooling, the steepness of the earnings profile and expenditures on education. This is fairly standard. Note that even though human capital investment decisions are efficient (bequests are positive in 2000 for almost all households), the presence of the spillover implies that the efficiency conditions are not exactly the same as the corresponding income maximization problem studied in Ben Porath (1967). The issue is that when an individual decides to invest in his human capital, he needs to take into account the effect it has on the human capital of his progeny and not just his own lifetime earnings. We need to calibrate the all-important spillover parameters γ_3 and γ_4 as well as the altruism parameters α_0 , α_1 and heterogeneity in z_h . We use aggregate moments in order to pin down these moments. Clearly micro evidence on the impact of parents on the human capital of their children can be brought to bear to pin down these spillover parameters - we postpone such a discussion to Section 6. The spillover parameter (especially γ_4) has a sizeable impact on

⁴The only difference between the life-cycle economy used to construct the initial condition and the economy presented here with zero bequests is that in the life-cycle model, the parent does not internalize the spillover. The quantitative effect is not that large for given values of real interest rates and wage rates since the individual does want to invest early on in the lifecycle.

the speed of transition and hence affects the level of output in 2000 relative to 1500 AD. The increase in GDP over the time period helps pin down this parameter. The altruism parameters affect the choice of fertility rate as well as bequests from parents to children. We use observations on heterogeneity of schooling levels in the cross-section to calibrate the distribution of z_h . We assume that z_h can take on 5 possible values. The moments we use to match α_0, α_1 and the values of z_h are in Table 1.⁵

Table 1: Calibration Targets

	Moment	Value	Source
1	GDP in 2000 relative to GDP in 1500	27	Maddison
2	Wage rate at age 55/wage rate at age 25	1.94	PSID
3	Years of schooling	12.45	Barro and Lee, 2010
4	Schooling expenditures/GDP	4.2%	UNESCO Institute for Statistics
5	Intergenerational transfers/GDP	4%	Gale and Scholz, 1994
6	Average Fertility Rate	2	World Bank data bank
7	Average Life span	80	World Bank data bank
8	Expenditures on health relative to GDP	10%	World Bank data bank
9	Average schooling level of top 10%	14.6	data from HRS
10	Average schooling level of top 70-90%	13.27	data from HRS
11	Average schooling level of bottom 10-30%	11.71	data from HRS
12	Average schooling level of bottom 10%	11.04	data from HRS

⁵A few comments are in order: All our targeted moments are for the United States with two exceptions. The increase in GDP per capita is for the UK. Second, health expenses relative to GDP in the United States are significantly higher than the corresponding figure for other developed countries and we use a number that is fairly typical in Europe.

Schooling is defined as the period of full time human capital acquisition ($n = 1$). Wages are assumed to reflect human capital less training costs. The view we take, which is fairly standard, is that when the individual of age i at time t gets trained at a firm, his training investments of time (n) and goods (x) inputs are subtracted from his human capital. Hence the observed wage equals $w_t h_{it}(1 - n_{it}) - x_{it}$. Wage growth over the life cycle reflects the fact that individuals spend less time on training as they age. While Mincer assumed a mechanical (linear) relationship between n and age, the relationship here is the result of an optimal choice. The model's counterpart to the wage growth is the average over all individuals in year 2000.

We obtain an exact match between model and data. The parameters are presented in the Table below.

Table 2: Model Parameters						
Parameter	z_h^1	z_h^2	z_h^3	z_h^4	z_h^5	μ
Value	0.165	0.167	0.171	0.179	0.183	0.89
Parameter	γ_1	γ_2	γ_3	γ_4	α_0	α_1
Value	0.60	0.30	0.07	0.83	0.69	0.63

Note that the returns to scale in the production of human capital, $\gamma = \gamma_1 + \gamma_2 + \gamma_3 = 0.97$ is high and the technology to produce human capital at age 6 also features a high returns to scale, $\gamma_4 = 0.83$. Unlike the endogenous growth literature, we use observations on schooling, earnings profiles as well as macroeconomic observations such as the increase in GDP per capita over a time period to pin down these parameters. The fact that the return to scale in the production of human capital is not *assumed* to be linear as in much of the endogenous growth literature but rather calibrated based on empirical counterparts is a virtue of our model. The effective returns to scale implied by our calibration is closer to 0.8 and in a subsequent section we report what happens when γ_4 is reduced. Moreover, we use micro evidence to argue that the calibrated values for the spillover parameters γ_3 and γ_4 are indeed reasonable when confronted with evidence on parent and child schooling. The calibrated degree of altruism suggests that while individuals care about their future generations, they place less weight on their children than their own utility. Furthermore, additional children bring in more utility but at a decreasing rate.

4 Results

We are now in a position to simulate the model's transition path from 1500 to 2000. We emphasize that there is no exogenous change along the transition. The initial human capital has to be sufficiently small in order to match a very low GDP per capita in 1500. At this low level, bequests are zero for all but the highest ability types. Most parents are poor enough that they would like to borrow against their childrens' income but are unable to do so. Hence, they under-invest in their childrens' human capital. There are two effects

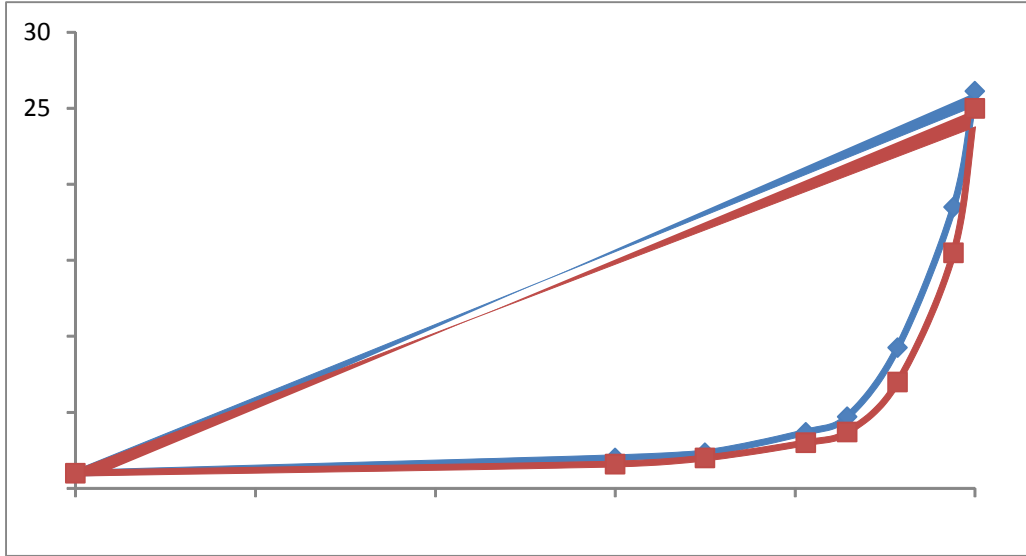


Figure 1: GDP per capita between 1500 and 2000

at play on fertility. First, there is a greater incentive to accumulate human capital after independence triggers. Second, higher incomes imply higher fertility. Over the course of time, human capital stocks rise and the marginal return to investing in human capital falls. At some point, dynasties get rich enough that they bequeath a positive amount. Human capital stocks and GDP rise rapidly from this point. The model is able to generate a long period of stagnation as well as rapid take-off after a while. Figures 2 (levels) and 3 (log-scale) show the reasonably close match between model and data for GDP per capita.

The model is able to generate the rise in years of schooling (Figure 4) consistent with the evidence as well as the decline in fertility (Figure 5).⁶ Indeed the ability to match the rapid take-off in GDP often referred to as the First Industrial Revolution as well as the dramatic fall in fertility, the Demographic Transition, without resorting to any exogenous force is one of the central results of this paper. As emphasized earlier, as fertility declines so does the real interest rate. Indeed this is a feature of the long time series that has led authors such as Alfred Marshall (1890, p. 403) to speculate that “Man, though still somewhat impatient of delay, has gradually become more willing to sacrifice ease or other

⁶There are two sources of non-monotonicity in the time series of fertility rates that the model is unable to capture: first, fertility rose before the demographic transition and second, fertility rose during the baby boom.

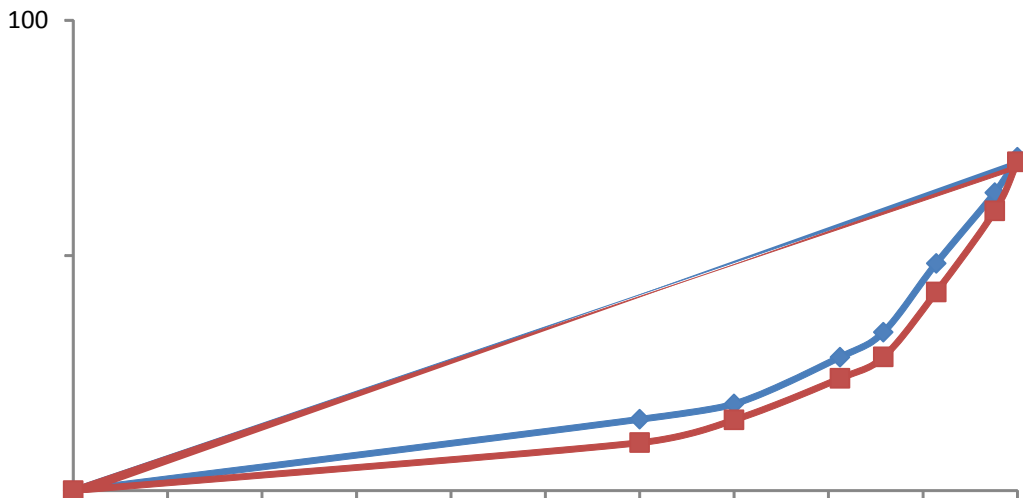


Figure 2: GDP per capita between 1500 and 2000 on log scale

enjoyment in order to obtain them in the future”. In our framework, parents affect their degree of patience towards the future members of their dynasty by choosing the number of children.

The relationship between real rates of return and fertility: The model can match the steep decline in fertility from 1850 and the relative constancy of real interest rates over the past 150 years. Given the calibrated parameter values, α_0 and α_1 , a small change in r implies with a fairly substantial change in f . Prior to 1850, the tight link between real interest rates and fertility rates does not hold since the bequest constraint binds. This subtle aspect of this dynastic model helps match time series observations of fertility and real rates of return and as we will see below, some cross-sectional aspects of the data. The real interest rate implied by the model falls from around 9% in 1500 to about 6% in 1800 and remains fairly flat after that, falling only very slightly to 5.5% as it reaches the steady state.

The Decline in Real Rates of Return: The decline in real interest rate plays an important role and hence it is useful to review the historical evidence. While we do not have one consistent series to measure real interest rates such as NIPA accounts, the real rate of return on riskless assets (one such asset was ‘rent charges’ which were paid over long periods of time) has declined over time. In England, rates of return on agricultural land

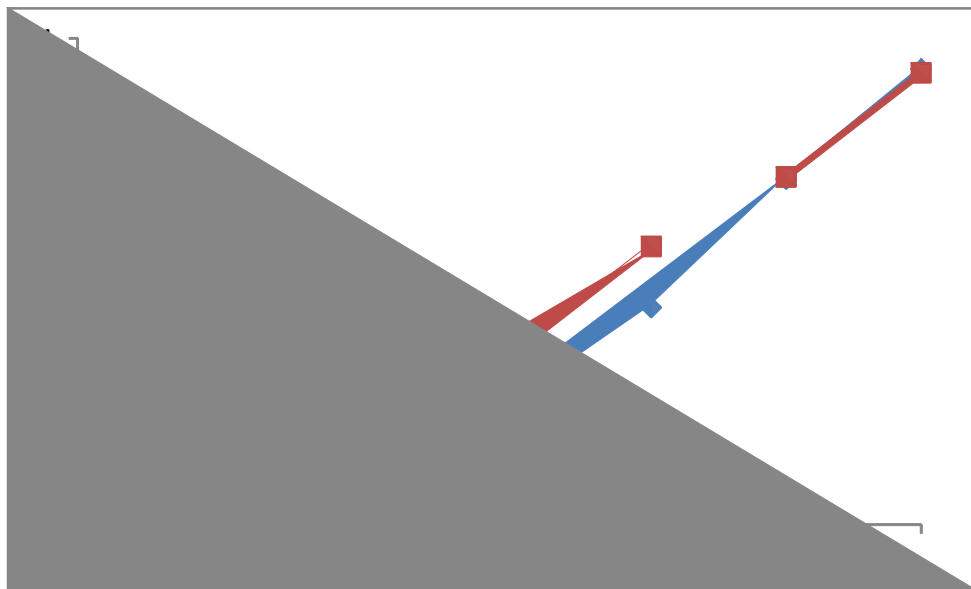


Figure 3: Years of Schooling: 1800 to 2000

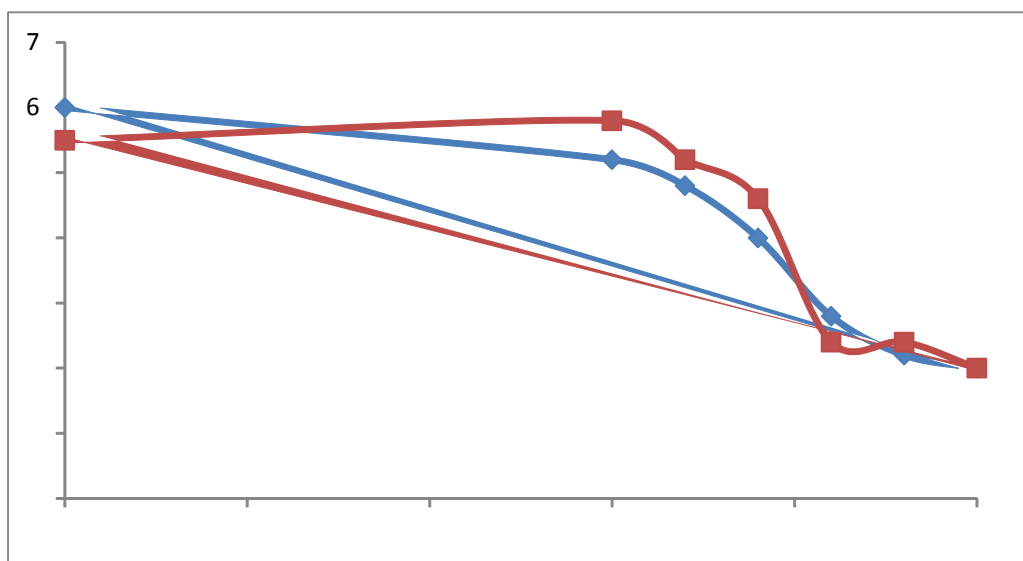


Figure 4: Total Fertility Rates: 1500 to 2000

were around 10 percent from 1150 to 1350, they fell to 5 percent and remained at that level until 1700 after which they fell to 3 percent. This decline was also seen in Northern Europe. For more evidence on this long term decline, see Clark (2005). Homer and Sylla's (2005) monumental work also documents the decline in real rates with a collection of data over long periods of time. Finally, Maddison's data indicate a considerable rise in the capital-output ratio over the time period under study which implies a fall in real interest rate via a standard Cobb-Douglas production function. The bulk of the evidence points to real rates declining until about 1800 after which they remained fairly constant.

4.1 Implications for Life Expectancy

As the developed economies took-off from a long period of stagnation, life expectancy began to rise as well. Life expectancy in England declined slightly from 1500 to 1700 after which it began to rise (data from Galor, 2005). Most of the increase happened towards the end of the 19th century and the rise was rather significant in the 20th century. How well does our model match this rise in life expectancy? The Figure demonstrates that it tracks this increase rather well. Until about 1800, there is not much of a change in lifespans. The majority of parents aren't yet at the point wherein they invest optimally in life extension. They are poor enough that their own marginal utility of consumption is rather high. Once they become rich enough, they invest in health as well their children's human capital. The resulting rise in life span also spurs human capital accumulation - a doubling of the lifespan has a significant effect on human capital decisions. The bottomline is that the model is able to rationalize this increase in lifespan even with a very parsimonious specification.

4.2 Changing Relationship between Fertility and Income

One remarkable feature of the transition that developed societies underwent is the changing relationship between fertility and economic status. A stylized fact in the modern era is the negative relationship between economic status and fertility. Indeed, it was precisely this cross-sectional feature that prompted Robert Barro and Gary Becker to reformulate the economic theory of fertility. Rather interestingly, this negative relationship did not always hold. Clark (2005) shows that in circa 1600 in England, completed fertility was

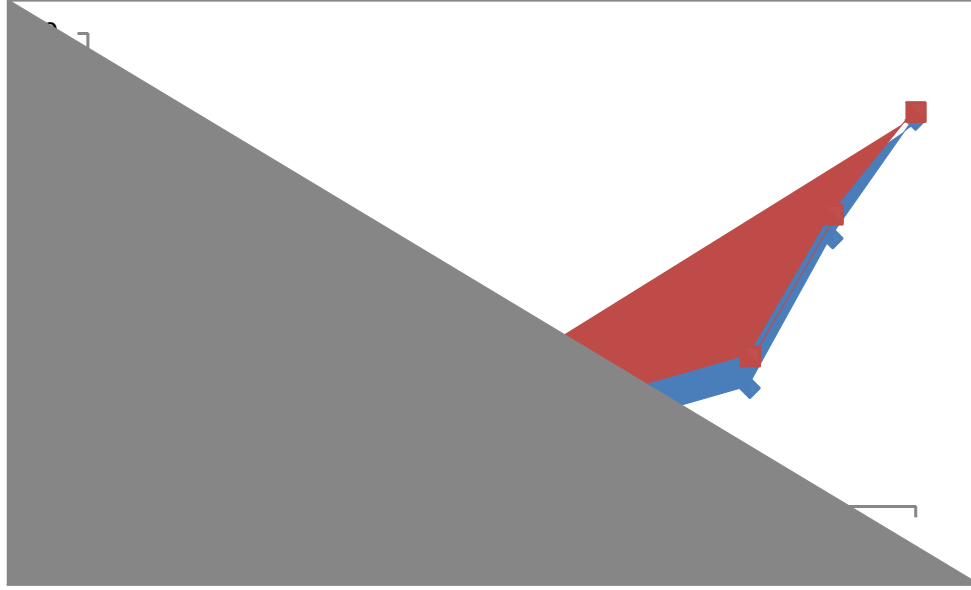


Figure 5: The Rise in Life Expectancy

positively associated with economic status - those who had larger estates chose to have more children. Similarly, Vogl (2013) demonstrates that in poorer countries in sub-Saharan Africa, durable goods ownership is positively linked to completed fertility.

Our model is consistent with this evidence. When societies are poor and human capital levels are low, parents do not find it optimal to spend much on their children's human capital. Further, poorer parents would ideally like to borrow against the future income of their children but are unable to do so. Children are on net costly leading poorer parents to have fewer children. As economies develop, richer parents spend significantly more on the human capital of their children (relative to poor parents). We note here that if time were the only input in the production of human capital, this result will not transpire. The incorporation of goods inputs x in the production of human capital results in the effective cost of having children to rise when economies develop. Consequently, fertility declines sooner for high ability families than for low ability families as the return to accumulating human capital for them is substantially higher. Low ability families face lower returns to human capital accumulation and hence experience a smaller increase in the net cost of having kids. This leads eventually to a negative relationship between fertility and learning ability.

4.3 The Decline in Concentration of Wealth

Before the industrial revolution, bequests were a significant part of non-human wealth. Bequests were likely a more important component of national wealth than even net investment. By contrast, bequests play a much less important role in modern times. Our model has stark predictions for bequests as well as the changing significance of bequests in wealth accumulation over long time horizons. While concrete evidence on the significance of bequests is not easy to come by, better information is available on the concentration of wealth. Clearly, there are various aspects of reality (most important of which may well have been primogeniture) that our model does not capture. Nevertheless, it is instructive to see if the mechanism in our model implies a concentration of wealth anywhere in the vicinity of what was observed. According to Lindert (1986), the top 10 percent of families in England and Wales owned around 90% of total (non-human) wealth between 1650 and 1900. Over the course of the 20th century, this number fell to around 50% in 1973. Our model's predictions for the same object go from 75% in 1500 to 40% in 2000. Moreover, our model implies that this fraction remains high until about 1880, which is consistent with the historical evidence. Over the course of the transition, poorer dynasties accumulate more human capital and become wealthier. As human capital levels rise, they first channel their resources to investments in human capital and subsequently get to the point where they are able to bequeath resources to their children. Before the industrial revolution, lifespans were short and the majority of society was poor. This meant the absence of either a lifecycle or an intergenerational motive to save for the vast majority of households. As time went by, lifespans rose and the economy became wealthier; both these motives took on more important roles. The importance of bequests in total wealth and the concentration of wealth declined. These are broadly consistent with the evidence.

4.4 The Effect of Initial Conditions on Development Paths

The analysis above suggests that the model economy has the potential to account for the path of development that many rich societies have undertaken. This path is featured by a long period of stagnation. As real interest rates and fertility rates begin to fall, the incentive to accumulate human capital rose. When societies became sufficiently well off,

growth accelerated. While the analysis thus far has shed light on the mechanism behind economic growth, it remained silent on the cause or the origin behind the growth process. A long line of work has investigated the origins of economic development. While there is considerable debate in the literature on the significance of various ‘fundamental’ causes, the focus here is on initial human capital. Glaeser et. al. (2004) and Barro (2012) argue that human capital leads to better institutions and that human capital is a more ‘basic’ source of growth than other factors. We are in a position to use our model to examine the implications of initial conditions on economic growth. Consider a society in which all households possessed half as much initial human capital in 1500 as in the baseline economy.

What happens to the long run path? We simulate the model to examine the implication of this change in initial condition. The results are striking: even after 500 years, the economies have about a 4 fold difference in GDP per capita. Convergence is slow and initial conditions have a very strong influence on development paths. It is very difficult to find good measures of human capital for circa 1500. While there are few direct measures of human capital that are available before the rise of formal schooling, two studies attempt to proxy for human capital and are worth mentioning. First, evidence from Baten and Zanden (2001) suggests that book production in 1450 (a proxy for human capital) is strongly related to long run development, consistent with the predictions of our theory. The second study uses the phenomenon of age-heaping that was pioneered by Joel Mokyr as a proxy for numeracy.⁷ Using this same method, A’Hearn et. al. (2009) find that Western Europe had reached high levels of numeracy and had diverged from the East before 1600.

By no means are we suggesting that this can offer a complete explanation of the development paths for different societies. That is a very difficult question and numerous historical events have altered the trajectories of different countries (as it would even in our model). Nevertheless, our model suggests that these differential paths may have something to do with initial levels of human capital. We have seen that initial conditions can have a rather large influence on when exactly the economy begins to take off. Once these

⁷Just as the ability to sign a document is considered a measure of literacy, one’s awareness of his or her own age is considered to be a measure of numeracy. Approximation of one’s age is manifested in a phenomenon referred to as age heaping.

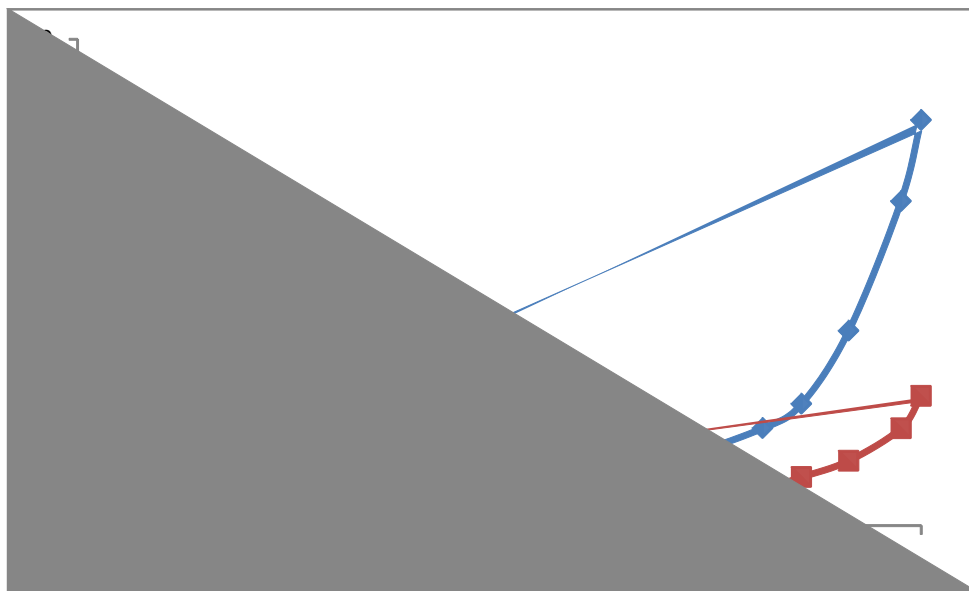


Figure 6: Impact of lower initial human capital on development paths

economies take-off, there is a tendency for convergence across economies. We now turn our attention to this.

5 Convergence

One aspect of the data that has been the subject of considerable scrutiny is whether economies converge starting from some initial condition and to what extent. Barro and Sala-i-Martin (1992) present an early analysis of convergence and show that the tendency of convergence implied by the Solow growth model is corroborated in the data. The primary focus of the empirical work has been cross-country data over the past 100 years. There have been numerous analyses that followed suit and different notions of convergence analyzed. In a recent paper, Barro (2012) concludes that the "iron law of convergence" holds in the data. Quoting Barro, "countries eliminate gaps of real per capita GDP at a rate of around 2% per year". At this rate of convergence, it takes around 115 years to eliminate 90% of the initial gap in per capita GDP.

We are in a position to use our model to gauge the ability of our model to rationalize the iron law of convergence. Consider the baseline model where the economy starts from

some initial condition. To conduct our analysis, it is advantageous to use long panels and the best available data lies in a recent data set put together by Jose Ursua and described in Ursua (2010). His dataset comprises of long panels of around 100 years of information on 42 countries. For each country, we obtain data on GDP per capita from 1870-2000 (the same time period analyzed in Barro, 2012). We keep the parameter values constant across countries and vary initial human capital by country to match GDP per capita at the first available year in which we observe GDP per capita in his sample. The starting dates vary considerably across countries ranging from 1790 for the United States to 1911 for Korea and South Africa. We then simulate each country until it reaches steady state. Using simulated data, we run the well known and widely studied regression for the time period 1870-2000,

$$\frac{1}{L} \log\left(\frac{Y_{it}}{Y_{io}}\right) = \alpha - \frac{1}{L}(1 - e^{-\beta L}) \log Y_{io} + \gamma X_{it} + u_{it}$$

where Y_{it} is GDP per capita in country i at time t , Y_{io} is GDP per capita when first observed (between 1790 and 1911), L is length of time for which growth is measured, X represents controls as is commonly done in growth regressions (here the controls are fertility rate, schooling and life expectancy) and β is the convergence coefficient. The estimated coefficient is $\hat{\beta} = 0.0176$ which means that it takes 130 years to eliminate 90% of gap. By way of comparison, the iron law of convergence at 2% per year implies that it takes around 115 years.

We conclude that the values of γ_3 and γ_4 result in rates of convergence that are fairly close to convergence rates estimated from long panels over the last 100+ years. The fact that the convergence rates are in the range of empirical estimates should not be a surprise to followers of the convergence debate. It is well known that standard capital shares imply convergence rates that much too fast. Barro and Sala-i-Martin (1992) show that in order to fit convergence patterns in the data, one needs a capital's share of around 0.8. Our analysis delivers exactly this. Recall that there are two technologies for human capital accumulation, one that is close to linear (with γ_3) and another determining human capital at age 6 which has a returns to scale of around 0.8 (γ_4). Given these two technologies, the effective returns to scale is around 0.8 (though some care needs to be exercised in comparing the finite horizon model we have presented here with the standard Neoclassical

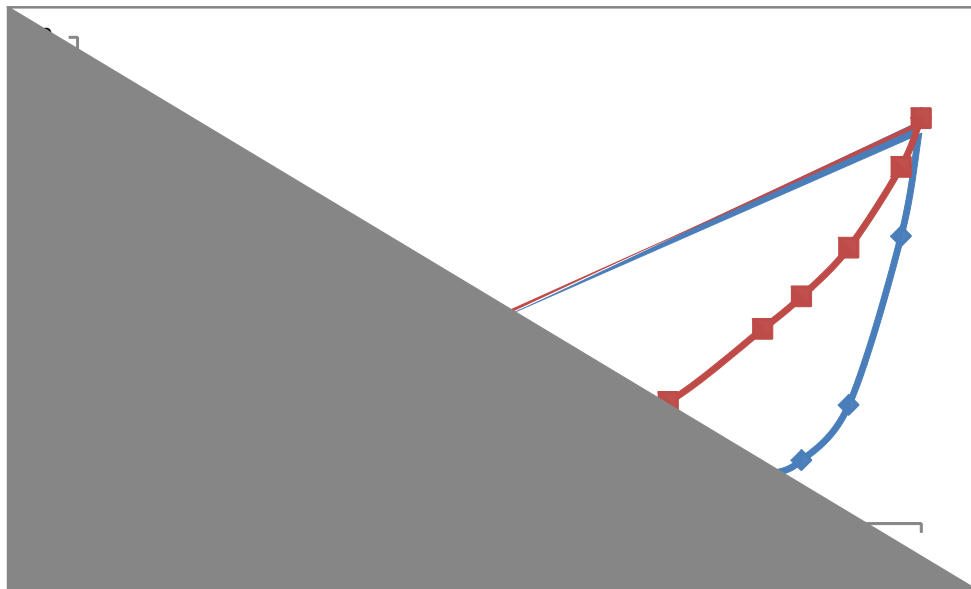


Figure 7: Effect of Lowering the Parental Spillover Term on the Transition Path

growth model where horizons are infinite). Our view is that a high value of the returns to scale is plausible when the concept of capital is enlarged to include human capital.

It is instructive to examine the impact of the spillover parameter γ_4 on convergence paths. Consider a thought experiment wherein γ_4 was cut in half. In addition, we adjust the initial stock of human capital so that we obtain the same increase in GDP per capita over the 500 year horizon. The results are in Figure 8. Holding everything else constant, simulations suggest a much faster convergence to the steady-state. Indeed, if γ_4 were zero, the model becomes an otherwise standard life-cycle model of human capital accumulation featuring relatively rapid convergence. This thought experiment suggests that the mechanism generating the reasonable time series predictions rests critically on the value of γ_4 .

We conclude that the available macro time series evidence lends support to our choice of parameter values. Specifically, the model can generate the long period of stagnation and subsequent growth and once rapid growth occurs, the model's implications for the speed of convergence over the past Century is in line with the evidence. Convergence in income per capita among OECD countries is well documented and our model is consistent with this empirical regularity. Nevertheless many poor countries do not grow and there is little

correlation between initial GDP per capita and growth rates for the world as a whole. Our model attributes the relative stagnation of poor countries to the fact that at very low levels of development, human capital accumulation decisions are inefficient. Indeed, the model predicts a much weaker correlation between growth rates and initial GDP among countries for which the bequest constraint is binding for a majority of households. We now proceed to examine the micro evidence to shed further light on the reasonableness of our parameter values.

5.1 Implications of Alternate Formulations

What would other formulations imply for the transition path? In this section, we look at the impact of some key assumptions on the long transition. While the analysis above demonstrates that our model economy is consistent with the iron law of convergence over the past 140 years, the previous section demonstrated that it can also generate this after a period of stagnation. The subtle combination of the life-cycle and the intergenerational elements along with the parental spillover term help in this regard. In each of the experiments below, we adjust the initial stock of human capital so that the starting point is essentially the same in terms of GDP per capita and then we examine the predictions of the model for convergence to the steady state. All the other parameters are held fixed at the same level as the baseline.⁸

Infinite horizons: Finite horizons play an important role. They force succeeding generations to accumulate human capital from an initial condition that depends on the human capital of their parents. As human capital levels rise, so does the initial stock of human capital. What if horizons were infinite? One way to accomplish this in our model (since lifespan is endogenous) is to make it costless to ‘produce’ lifespan, that is to set $\mu = \bar{T} = \infty$. What happens when $\mu = \bar{T} = \infty$? The result is a fairly standard infinite horizon model featuring significantly more rapid convergence. Starting from an initial stock of human capital, infinitely-lived households accumulate human capital at a rapid

⁸A more stringent test would be to fit both the initial condition and the end point. It turns out that in most cases, such parameters do not exist. This suggests that these alternative formulations need something exogenous changing over time in order to explain the more than 25 fold rise in GDP per capita over the period in question.

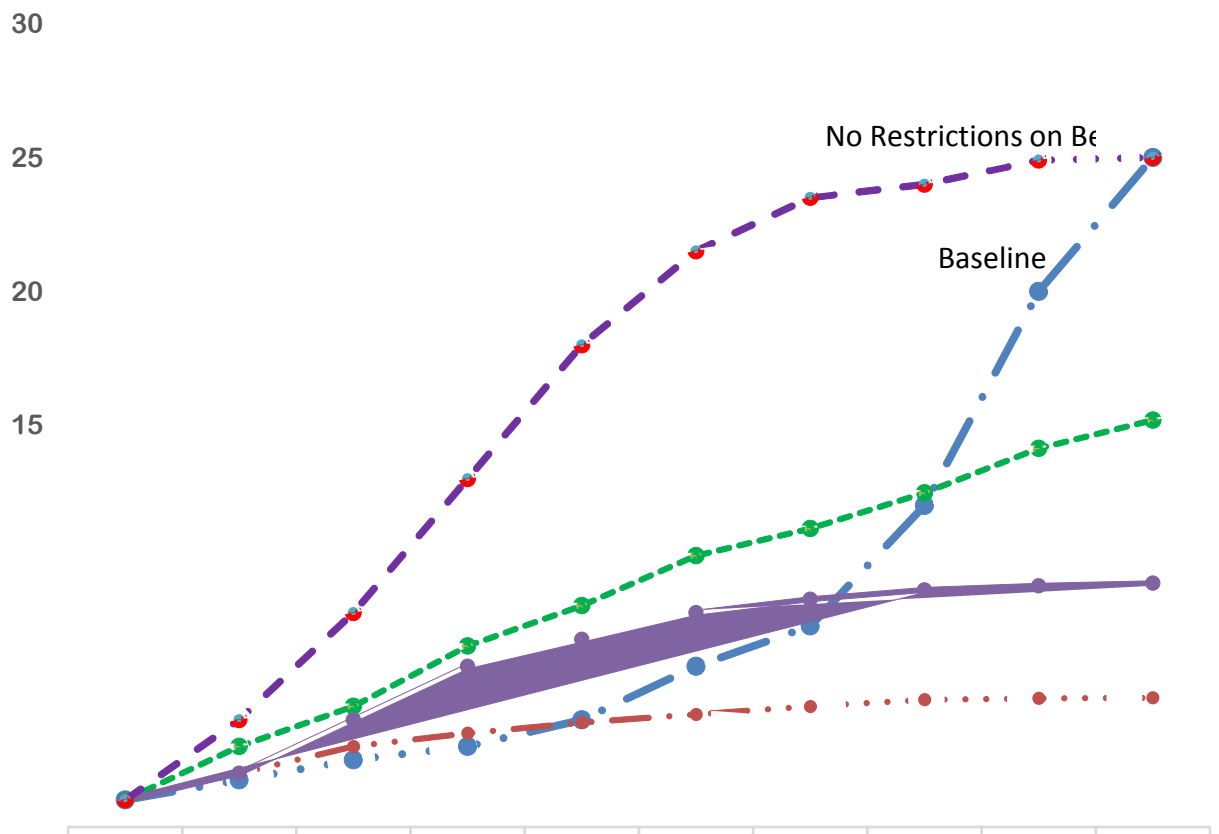


Figure 8: Implications of Alternative Assumptions for Convergence

pace and the speed of convergence mimics that in Barro and Sala-i-Martin (1992) when the capital's share of GDP is set at a number closer to 0.8. In the finite horizon version of our model, households need to accumulate human capital from scratch whereas in the infinitely lived households decision problem, no such parallel exists.

Endogenous Growth: A standard formulation in economics is to have infinite horizons and endogenous growth. This can be accomplished by having $\mu = \bar{T} = \infty$ and all the human capital production technologies linear. What happens when $\gamma_4 = 1$ and $\gamma_1 + \gamma_2 + \gamma_3 = 1$? As is well known, when we have linearity as well as infinite horizons, regardless of where the economy begins, it 'jumps' to the balanced growth path right away and hence features instantaneous convergence. The resulting framework can account for the constancy in growth rates (endogenously) but cannot account for the transition. This model does imply that growth is at a constant value forever.

No *Altruism:* Altruistic parents invest in their own human capital not just because they care about their lifetime earnings, but also because their human capital spillover to their children. Furthermore, after a certain point in time, they bequeath resources to their children and this has the effect of increasing the stock of capital and decreasing the interest rate. Clearly, α_0 has an important effect on the speed of convergence. What happens when $\alpha_0 = 0$? Altruism leads to a higher stock of physical capital (through bequests) and human capital (since parents invest more in themselves taking into account the effect on future generations). The economy without any altruism needs a higher initial stock of human capital in order to match the same GDP per capita as the baseline figure. The model features a period of stagnation that is shorter and ends up converging to a lower level of GDP per capita in the long run. Lower degrees of altruism are associated with higher real interest rates which hinder human capital investments.

No Restriction on Bequests: The critical element generating the slow transition to the steady state is that initially bequests are zero. The non-negativity constraint on bequests binds. Hence, for a long period of time, generations are unconnected through bequests and invest inefficiently in their human capital. After a certain point in time, dynasties get rich enough that the vast majority bequeath resources to their children and invest efficiently in them. After that point in time, convergence is rather rapid. *The Changing nature of dynastic linkages* plays a critical role in generating a path that resembles the data. What

happens if b were unconstrained? Even poor dynasties in 1500 can now borrow against the future incomes of their dynasties. Real rates of return converge more rapidly to their steady state values and so do human capital stocks and GDP.

In the baseline economy, the rich are unconstrained in circa 1500. As they get richer and bequeath more to their children, real interest rates fall as the capital stock rises.⁹ The reason for this is that the stock of physical capital rises more rapidly than the stock of human capital. Physical capital rises primarily because of increased bequests by the top 10% of the households. Given that lifespans did not change much between 1500 and 1800 (see Figure 6), the life-cycle motive for saving did not play a dominant role. The fall in the real interest rate induce other households to invest greater amounts in human capital. Eventually, most of the households leave positive bequests (due to this trickle-down General Equilibrium effect) and invest efficiently in the human capital of their children. If there are no restrictions on bequests, all households will be investing efficiently from the get go. Consequently the path of GDP per capita from the initial condition mimics the pattern in the baseline economy from 1820 featuring rather flat real interest rates since both physical and human capital rise by roughly the same amount. The non-negative restriction on bequests thus plays an important role. It interacts with heterogeneity in a fairly subtle manner contributing to the period of stagnation as long as the constraint binds for a majority of households and also contributing to the eventual take-off when most households begin investing efficiently in their children.

6 Micro Evidence on the Spillover

One of the central advantages of our framework relative to say one with aggregate spillovers is that it is possible to measure the parental spillover by using individuals' earnings and schooling data. In this section, we use our theory to generate clear predictions at the

⁹King and Rebelo (1993) show that neoclassical model implies rapid transitions unless marginal products of capital and interest rate are extraordinarily high in the early stages of development or the capital share is much higher than observed in national accounts. Our model succeeds in explaining a lengthy transitions without these counterfactual implications. The capital share is constant along the transition and interest rates are matched up with historical evidence.

micro level that corroborate the calibrated spillover values. A robust finding in empirical studies is that even after controlling for observables, a mother’s education has a significant impact on children’s schooling and earnings. In many ways, our intergenerational spillover can be thought of as capturing just this. The question is whether the magnitude of the effect matches up with the evidence. This section explores this issue in greater detail and in particular sheds light on the important issue of how to use micro evidence to estimate the spillover parameters: γ_3 and γ_4 .

6.1 Empirical Evidence

We examine data from the 1926-1941 cohort of the Health and Retirement Survey (HRS) for which complete earnings histories are available. A data appendix describes the data as well as the sample in detail. Figure 9 presents experience-earnings profile for male high school graduates controlling for cohort effects. We split the individuals into two subsamples depending on whether or not the mother has more than 6 years of schooling. The estimated profiles support the importance of parental human capital since we observe higher earnings for individuals with more educated mothers. It appears essentially as a permanent level effect that persists throughout an individual’s career with no evidence of increasing steepness of age-earnings profile.

We now examine the role of parental human capital on the age-earnings profiles of their children to infer the implied values of the parental spillovers in the human capital production function. We augment a standard Mincer regression with measures of parental human capital. We estimate the regression

$$\log w_i = \alpha_0 + \alpha_1 S_i + \alpha_2 S_i^p + f(E_i) + \beta X_i + \epsilon_i,$$

where w_i , E_i , S_i and S_i^p denote, respectively, earnings, potential experience, years of education of individual i , and years of education of individual i ’s parents. f is a flexible function of potential experience, X_i is a vector of demographic control variables¹⁰ and ϵ_i is an error term.

Traditionally, the human capital of an individual has been approximated by $h_p = \exp(S^p)$, where S^p is the schooling level of the parent. This follows a standard Mincerian

¹⁰We control for race and cohort effects.

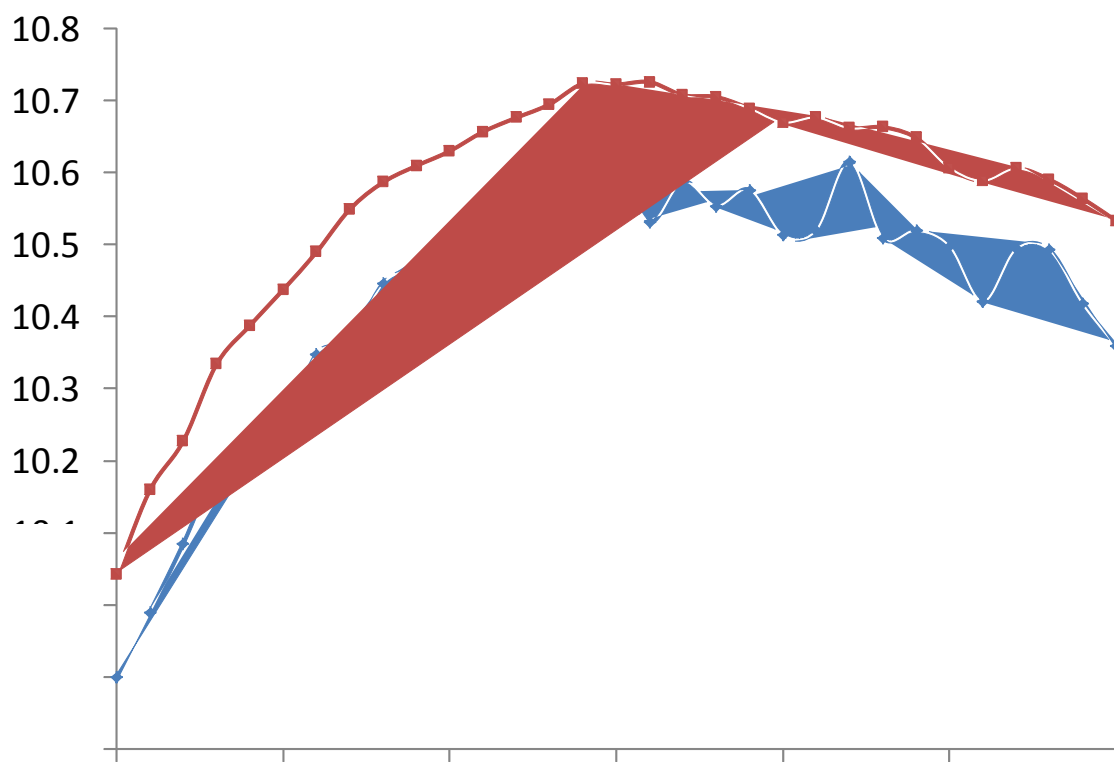


Figure 9: Earnings Profiles of High School Graduates by Mother's Schooling Level: 1926-1941 Birth Cohort

equation. In our framework this relationship is an equilibrium outcome and this log-linear relationship is indeed a good approximation of the optimal solution. Indeed, a good approximation of the Ben-Porath model presented above which features the quality of human capital is $h_p = \beta_0 \exp(S_i^p)$, where β_0 is the intercept term in the Mincerian wage regression. We consider three measures of parental human capital: mother's years of education, father's years of education, the sum of both parent years of schooling. Our theory is silent on which of these measures is more appropriate. Yet, in fitting the historical patterns of GDP, we concluded that parental human capital has its strongest impact on early childhood human capital (the γ_4 effect). It is commonly argued that mothers spend more time with their children at an early age which suggests that mothers' education should play a dominant role.

We estimate different specifications of this equation using data from the HRS. We restrict the sample to men born between 1926-1941 and aged between 25 and 55. Table 3 shows the multivariate regression results of log earnings on schooling, schooling of the parent, cohort and race dummies. The first specification (1) is a standard Mincer regression and it uses dummies for each potential experience level observed in the data. The return to schooling is estimated to be 6.57%. It is in the lower range of the estimated returns to schooling with more recent cohorts which is in line with the increased return to education over the last century (see Goldin and Katz, 2008). This returns slightly decreases to 5.99% when we include mother's years of education in the regression (column 2). The coefficient on mother's education is 1.27% and is statistically significant. Our estimates imply that an additional year of schooling for the child has about the same effect on earnings as would being born to a mother with five additional years of schooling. The results are quite similar when we measure parental human capital with the father's year of education (column 3) but with an attenuated coefficient. The rate of return of paternal education is 0.92% and is statistically significant. The coefficient drops further when we measure parental human capital with the sum of the schooling of both parents (column 4). If we include both (column 5), mother's education is very slightly reduced from 1.27% to 1.16% while the coefficient on fathers drop significantly from 0.93% to 0.26%.

We also experimented with two interactions terms: between parental education and experience and between education and experience and where we replaced the experience

dummies with a quadratic in experience (column 6). We find that the interaction between parental education and experience is insignificant. Finally, the results are quite stable when the sample is restricted to white males only (column 7).

Overall, we find that mothers have a stronger impact on sons than fathers and that the effect of parental education is about 1%. These results are in line with those reported by previous empirical work on this topic (see for instance Card, 1999 and references therein). In light of this evidence, we can now go back to the model and ask what they imply for the values of γ_3 and γ_4 .

6.2 Implications for Parental Spillovers

In the model economy presented above, ability z_h has a positive influence on schooling. The effect of parental human capital h_p is slightly more complex. Parental human capital serves two roles: first, a higher parental human capital level implies a higher initial stock of human capital for the child, which all else equal, decreases schooling. Second, a higher parental human capital level effectively increases the ability to learn for the child and this effect counters the first effect.

Our model has implications which we have verified for a broad range of parameter values using model simulations for the innate abilities of two individuals with the same years of schooling but whose parents have different schooling levels. Let us define two terms for ease of exposition. Denote effective ability $a = z_h h_p^{\gamma_3}$ and $h_p^{\gamma_4}$ effective initial human capital. As long as $\gamma_4 > 0$, a child of a higher human capital parent will start off with more effective human capital and earn more conditional on schooling. When γ_4 equals zero, these two individuals with the same schooling level start off with the exact same effective human capital and have similar earnings profiles in our simulations. This is by virtue of the fact that the individual with the higher human capital parent must necessarily possess a lower innate ability, z_h so that they have the same effective ability.

Hence looking at earnings differences between two individuals who possess the same schooling level but have parents with different human capital levels identifies γ_4 . If there is no earnings differential whatsoever across these individuals with the same years of schooling but different parental human capital levels, the level effect governed by γ_4 must

Table 3: Augmented Mincer Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Education	0.0657*** (113.27)	0.0599*** (86.57)	0.0609*** (86.44)	0.0592*** (79.90)	0.0591*** (79.87)	0.0869*** (46.15)	0.0600*** (78.15)
Mother Ed		0.0127*** (21.63)			0.0116*** (15.55)	0.0157*** (10.83)	0.0130*** (20.12)
Father Ed			0.0092*** (17.03)		0.0026*** (3.93)		
Parental Ed				0.0068*** (20.73)			
Exp.						0.116*** (64.90)	
Exp. ²						-0.0018*** (-74.38)	
Exp. \times Ed.						-0.0013*** (-17.34)	
Exp. \times Mother Ed.						-0.0001 (-1.95)	
Observations	171194	150684	146077	140380	140380	150684	127666
R^2	0.156	0.154	0.151	0.151	0.152	0.152	0.133

OLS regressions of (log) earnings on years of schooling and years of schooling of the parents controlling for potential experience, race and cohort effects. The Table reports the number of observations and the R^2 . t-stats are shown in parentheses. Age 25-55, cohort born 1926-1941, Male, earnings below first percentile are trimmed.

equal 0. On the other hand, sizeable lifetime earnings differentials between these two individuals will be associated with large values for γ_4 .

The estimates suggest that an extra year of parental schooling is associated with 1% higher wages for the child, *conditional* on the schooling of the child. We divide the data into 10 deciles by lifetime earnings (conditional on 12 years of schooling). For each of these deciles, we calculate the mean maternal schooling level. This gives us an estimate of h_p . We then vary z_h so that the optimal schooling level is precisely 12. We then run the same regression in our model. Clearly, when γ_4 equals zero, this estimate of return to maternal schooling conditional on the schooling level of the child is zero. We ask what value of γ_4 allows us to obtain a 1% estimate. The estimate of γ_4 is 0.77. Note that this estimate is independent of the value of γ_3 in the sense that we implemented this procedure with different values of γ_3 and the resulting estimate of γ_4 is not sensitive to the values of γ_3 .

How do we identify γ_3 from the data? We proceed to identify γ_3 by looking at two individuals with different years of schooling and examining their wages later on in life. Given the estimate of γ_4 we look at data on individuals with 12 and 16 years of schooling. Those with 12 years of schooling had mothers with average years of schooling of 9.35 while those with 16 years of schooling had mothers with average schooling of 11. Given the estimate of γ_4 above, we know the impact that this maternal schooling difference has on initial human capital at the age of 6. The difference in log earnings between college and high school grad at age 45 is 0.3. We ask what value of γ_3 can precisely rationalize the college-high school earnings differences when innate ability z_h is varied to rationalize the fact that one group of children acquire 4 additional years of schooling relative to the high school group. The estimate of γ_3 is 0.05.

The exercise above suggests a novel way of estimating intergenerational parental human capital spillovers. The model has stark implications for the impact of parental schooling on children's outcomes and by conditioning on schooling of the child, we are able to get around some of the selection issues commonly encountered in empirical work. There is a body of work estimating the intergenerational correlation in schooling across generations, which is typically estimated to be in the 0.25 to 0.4 range. More recently, there has been interest in uncovering the causal relationship between parental schooling and children's wages by using compulsory schooling laws as an instrument. This body of work has argued that

that there is little discernible causal effect of parents education on childrens' education. They use compulsory schooling laws to argue that parents who acquired more schooling than they had planned on (due to a shift in government policy) did not have children with appreciably more schooling. On the face of it this calls into question the existence of substantial spillovers from parent to child human capital. Remarkably, our model is consistent with the rather low IV estimates of parental schooling on children's schooling and this is something we explore in greater depth in ongoing work. We conclude that micro evidence lends some support to our specific parental spillover terms.

Finally, it is useful to mention the significance of nature and nurture. For the model presented above, the human capital spillover is a learning effect which we consider the model's counterpart to nurture. One could think of a more general model with noise (by assuming that ability is not perfectly transmitted across members of the dynasty but follows an AR(1) process) in the transmission of z_h across generations. This would be the counterpart to nature. The analysis presented above demonstrates a way to estimate BOTH γ_3 and γ_4 without having to deal with the nature-nurture decomposition, which is a difficult problem. In on-going work we use our model to shed light on this decomposition. Depending on the relative importance of γ_3 and γ_4 , our model has different implications for the innate abilities of individuals that went to school for the same number of years but have parents with different years of schooling. If the γ_3 effect dominates, a child of a higher human capital parent will have a lower innate ability z_h than his counterpart whose parent possess lower human capital. If the γ_4 effect dominates, a child of a higher human capital parent will possess higher innate ability than his counterpart whose parent possesses lower human capital. Since innate ability is likely positively correlated across generations (nature), the latter case is the empirically relevant case.

7 Conclusion

Two of the most commonly used models in the study of economic growth are the exogenous growth and the endogenous growth models. The unsatisfactory aspect of the former lies in its heavy reliance on an exogenous variable, TFP while the unpalatable feature of the latter is that the linearity assumption is rather ad-hoc. To be clear, we are not suggesting that

technological progress is unimportant but rather that understanding growth by relying so heavily on an exogenous variable (TFP) is simply satisfactory. This sentiment is best expressed by Arrow (1962) who explores the economic implications of learning by doing.

It is by now incontrovertible that increases in per capita income cannot be explained simply by increases in the capital labor ratio. Though doubtless no economist would ever have denied the role of technological change in economic growth, its overwhelming importance relative to capital formation has perhaps only been fully realized with the important studies of Abramowitz and Solow. The results do not contradict the neo-classical view of the production function as an expression of technological knowledge. All that has to be added is the obvious fact that knowledge has been growing in time. Nevertheless a view of economic growth that depends so heavily on an exogenous variable, let alone one that is so difficult to measure as the quantity of knowledge, is hardly intellectually satisfying.

Kenneth Arrow (1962)

In this paper, we take this challenge of rationalizing growth, without resorting to dependence on an exogenous variable or ad-hoc assumptions of linearity, seriously. We attempt to provide a model with which one can quantify knowledge in an empirically verifiable fashion. We depart from the well known framework in Lucas (1988) in three very important respects: first, horizons are finite, a feature that helps generate empirical counterparts; second, the technology to produce human capital features diminishing returns (a property that arises by forcing the model to match observations) and third and most important, we move away from the economy-wide spillover and instead model a parental spillover.

We present a model along the lines of the pioneering work of Barro and Becker (1989) that can rationalize the fall in real interest rates. As is well known, models of fertility choice feature a relationship between real interest rates and fertility rates. From an initial condition in 1500, we simulate the transition to a steady state. The presence of a parental spillover is critical to generating a transition that is consistent with the evidence. The model generates a path of GDP per capita that is not too dissimilar to the observed paths followed by Western European countries. We show that if the spillover term is sufficiently

large, the model can generate stagnant living standards until 1800 and rapid growth from that point on. Moreover, the dramatic rise in living standards is accompanied by fertility declines and schooling increases that match up with the data fairly well. We also show that initial levels of human capital can have a rather large influence on development paths. Economies that began with a lower stock of human capital in 1500 develop almost a century later. We draw the conclusion that human capital can serve as both an *origin* as well as a *cause* of economic growth.

The parental spillover term plays a critical role in determining the speed of transition. Clearly, being able to estimate the size of the spillover term based on macro and micro evidence is critical to justifying the mechanism at work. Using the model's implications for a panel of more than 40 countries for the past 120 years, we show that the calibrated value of the parental spillover term is consistent with the iron law of convergence. We then demonstrate that our model has sharp implications at the micro level. Consider two children who possess the same years of schooling but have different parental human capital levels. Children with more highly educated parents earn substantially more than their counterparts who are raised by less educated parents even after controlling for schooling levels. The level differences in log earnings between these two children sheds light on the spillover term and we show that the estimated value is close to the calibrated value that is picked to match long time series observations. We conclude based on the evidence that the key mechanism at work has empirical support.

References

- [1] Acemoglu, D., Johnson, S., and Robinson, J. A., 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation". *American Economic Review* 91(5): 1369-1401.
- [2] Aghion, P., Howitt, P.W. (1992). "A model of growth through creative destruction" *Econometrica* 60, 323-351.
- [3] A'Hearn, Brian & Baten, Jörg & Crayen, Dorothee, 2009. "Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital," *The Journal of Economic History*, Cambridge University Press, vol. 69(03), pages 783-808, September.
- [4] Arrow, Kenneth J., "The Economic Implications of Learning by Doing" 1962. *The Review of Economic Studies*, Vol. 29, Issue 3, p. 155-173 1962
- [5] Barro, Robert J., 1997, Determinants of Economic Growth. Cambridge, MA: MIT Press.
- [6] Barro, Robert J., 2012, "Convergence and Modernization Revisited," NBER Working Paper No. 18295.
- [7] Barro, Robert J and Gary S. Becker, 1989, "Fertility Choice in a Model of Economic Growth," *Econometrica* 57 (2): 481–501.
- [8] Barro, Robert J and Jong-Wha Lee, April 2010, "A New Data Set of Educational Attainment in the World, 1950-2010." NBER Working Paper No. 15902.
- [9] Barro, R. J., and X. Sala-i-Martin. 1992. "Convergence." *Journal of Political Economy* 100 (April): 223–51.
- [10] Baten Joerg & Jan Zanden, 2008. "Book production and the onset of modern economic growth," *Journal of Economic Growth*, Springer, vol. 13(3), pages 217-235, September.

- [11] Becker, Gary S & Murphy, Kevin M & Tamura, Robert, 1990. "Human Capital, Fertility, and Economic Growth," *Journal of Political Economy*, University of Chicago Press, vol. 98(5), pages S12-37, October.
- [12] Becker, Gary S. and Tomes, Nigel, (1986) "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics* 4: S 1 -S39.
- [13] Ben Porath, Y., 1967, "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy* 75, pt. 1, 352-365.
- [14] Browning, M., L. P. Hansen and J. J. Heckman, 1999, "Micro Data and General Equilibrium Models," in J. B. Taylor and M. Woodford (eds), **Handbook of Macroeconomics**, Vol. 1, North Holland, pp: 544-633.
- [15] Card, D. (1999). The causal effect of education on earnings. *Handbook of labor economics*, 3, 1801-1863.
- [16] Cass, D. (1965). "Optimum growth in an aggregate model of capital accumulation". *Review of Economic Studies* 32, 233-240.
- [17] Clark, Gregory, 2005 "The Interest Rate in the Very Long Run: Institutions, Preferences and Modern Growth," Working Paper, UC-Davis.
- [18] Engerman, S. L. and Sokoloff, K. L. (1997). "Factor Endowments, Institutions, and Differential Paths of Growth Among New World Economies: A View from Economic Historians of the United States." In *How Latin America Fell Behind*, edited by Stephen Haber, 260–304. Stanford, CA: Stanford University Press.
- [19] Galor, O., 2005, "From Stagnation to Growth: Unified Growth Theory," in P. Aghion and S. Durlauf (eds) **Handbook of Economic Growth**, North Holland.
- [20] Goldin, C. and Katz, L. *The Race Between Education and Technology*. Belknap Press for Harvard University Press; 2008.
- [21] Glaeser, E. L., La Porta, R., Lopez-de-Silanes, F. and Shleifer A. (2004). "Do Institutions Cause Growth?" *Journal of Economic Growth*, 9: 271-303.

- [22] Grossman, G.M., Helpman, E. (1991). *Innovation and Growth in the Global Economy*. MIT Press, Cambridge.
- [23] Heckman, J., L. Lochner, and C. Taber, 1998, "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics* 1, 1–58.
- [24] Homer, Sidney and Richard Sylla. 2005. **A History of Interest Rates**. Fourth Ed. New York: Wiley.
- [25] Jones, Charles, 1995 "Time Series Tests of Endogenous Growth Models," *Quarterly Journal of Economics*, May 1995, Vol. 110, pp. 495-525.
- [26] Koopmans, T.C. (1965). "On the concept of optimal economic growth. In: *The Economic Approach to Development Planning*. North-Holland, Amsterdam.
- [27] Lindert, Peter. 1986. "Unequal English Wealth since 1670." *Journal of Political Economy* 94 (6): 1127-62.
- [28] Lucas, R., 1988, "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22, 3-42
- [29] Maddison, Angus (2003). *The World Economy. Historical Statistics*. Paris, OECD. Maddison,
- [30] Marshall A (1890). **Principles of Economics**, 8th ed. Macmillan, London
- [31] Mankiw, N Gregory, Romer, David & Weil, David N, 1992. "A Contribution to the Empirics of Economic Growth," *The Quarterly Journal of Economics*, MIT Press, vol. 107(2), pages 407-37, May.
- [32] Manuelli, Rodolfo E. & Ananth Seshadri, 2009. "Explaining International Fertility Differences," *The Quarterly Journal of Economics*, MIT Press, vol. 124(2), pages 771-807, May.
- [33] Mincer, J., 1974, **Schooling, Experience and Earnings**, Columbia University Press, New York, 1974.

- [34] North, D.C., Thomas, R.P. (1973). The Rise of the Western World: A New Economic History. Cambridge University Press, Cambridge, UK.
- [35] OECD, 1996, **Education at a Glance: OECD Indicators**.
- [36] North, D.C., Thomas, R.P. (1973). **The Rise of the Western World: A New Economic History**. Cambridge University Press, Cambridge, UK.
- [37] Sachs, J.D. (2001). "Tropical Underdevelopment". NBER Working Paper No. 8119
- [38] Solow, Robert M., 1956, "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, Vol.70 No.1, pp: 65 - 94.
- [39] UNESCO Institute for Statistics, Global Education Digest 2007.
- [40] Ursúa, José F. (2010) "Macroeconomic Archaeology: Unearthing Crises and Long-Run Trends," Harvard University, unpublished.
- [41] Vogl, Tom, 2013 "Differential Fertility, Human Capital, and Development," Working Paper, Department of Economics, Princeton University.

8 Data Appendix

The Health and Retirement Study (HRS) is sponsored by the National Institute of Aging and conducted by the University of Michigan with supplemental support from the Social Security Administration. The HRS is a national panel study with a sample (in 1992) of 12,652 persons in 7,702 households. It oversamples blacks, Hispanics, and residents of Florida. The sample is nationally representative of the American population 50 years old and above. The baseline 1992 study consisted of in-home, face-to-face interviews of the 1931–41 birth cohort and their spouses, if they were married. Follow up interviews have continued every two years after 1992. As the HRS has matured, new cohorts have been added.

For this paper, we start with 30,548 individuals in the RAND HRS Version J (RAND, June 2010). We keep 6967 male respondents born between 1924 and 1941. We drop 32

individuals with missing information on years of schooling. Another 127 individuals are dropped because we could not obtain any of their earnings histories between age 25 and age 55 even with the above model. In the paper we measure earnings growth between age 25 and age 50 as the ratio of average earnings between ages 49 and 51 relative to the average earnings between ages 24 and 26. We thus drop another 1131 observations for whom we have no earnings data either between 24 and 26 or between 49 and 51. This leaves us with 5677 observations.