Author: Peter Galbács

Affiliation 1: Associate professor of economics, Budapest University of Economics and Business.

Affiliation 2: Rosztoczy Foundation visiting researcher, Kenneth C. Griffin Department of Economics, the University of Chicago.

Title: Rational drive for school segregation in a model of learning.

Funding details: This work was supported by the Rosztoczy Foundation.

Disclosure statement: The author reports there are no competing interests to declare.

Word count: 17352.

Rational drive for school segregation in a model of learning

Peter Galbács[♣]

January 31, 2026.

Abstract: We develop a dynamic two-sector model of formal schooling where student human capital evolves with study time, prior knowledge, and positive peer spillovers à la Lucas (1988). Families observe classmates' innate learning efficiencies and self-segregate to maximize their child's skill trajectory. We obtain closed-form steady states and show saddle-path stability for decentralized equilibrium and social planner's optimum. Results reveal asymmetric sorting: high-ability households strongly prefer similar peers, while low-ability households exhibit weaker segregation incentives. By locating knowledge spillovers within schooling, the model isolates a mechanism driving socioeconomic clustering. We discuss abstractions and propose extensions for information frictions and policy analysis.

Keywords: human capital accumulation, optimal control, segregation, peer effects, externalities

JEL codes:

[♣] Associate professor of economics, Budapest University of Economics and Business. Email: Galbacs.Peter@uni-bge.hu

1. Introduction

Lucas (1988) introduced knowledge spillovers in production as the engine of endogenous growth, framing knowledge as a factor whose productivity depends on the average knowledge of society. Empirical work reveals a parallel phenomenon in classrooms—peer-effects—where classmates' knowledge (or simply the quality of the learning environment) boosts the efficiency of individual human-capital accumulation. Anticipating these peer-driven gains, families strategically choose schools for their children, helping to explain why classrooms often cluster along socioeconomic lines.

Extensive research documents the adverse consequences of school segregation on student outcomes. Schools sorted by race or income typically suffer from inadequate funding, fewer experienced teachers, and limited instructional resources—conditions that depress test scores and long-run prospects (Borman, Hewes, Overman, & Brown, 2003; Weathers & Sosina, 2022). Recent studies also link segregation to setbacks in cognitive development and psychological well-being among disadvantaged students (Gutierrez, et al., 2024; Peterson, et al., 2021). In this account, segregation stands as the root cause of material and institutional barriers that erode learning.

Since the Coleman Report (Coleman, et al., 1966), a related strand of literature has highlighted peer-effects as a key channel through which students influence one another's outcomes (Sacerdote, 2011). Classmates' socioeconomic backgrounds and achievement levels shape individual test scores, motivation, and long-term success. Segregation magnifies these dynamics by grouping students with similar challenges, which can limit the academic stimulation and support found in more integrated settings. Empirical and theoretical work shows that expectations about peer quality drive families' school-choice decisions, reinforcing segregation as households seek the most advantageous learning environment (Manski, 1993; Epple & Romano, 1998; Orfield & Lee, 2007). From this perspective, socioeconomic segregation stems not only from disparities in funding and institutional rules but also from families' rational responses to anticipated peer-effects.

In this paper, we isolate why rational households desire to self-segregate by learning ability, without modeling segregation's full emergence via search or matching. We embed peer-effects into a two-sector Uzawa–Lucas framework, relocating Lucas's (1988) externality from production to the learning process. Under our benchmark assumption, families perfectly observe prospective classmates' innate learning efficiencies—a clean setup that yields closed-form steady states and transparent comparative statics. In Sector 1, human-capital accumulation is governed by a learning function that takes study time, previously acquired knowledge, and a positive peer externality as inputs. Sector 2 then translates accumulated human capital into utility through a standard consumption–leisure trade-off.

We present this model as a positive, mechanism-focused benchmark—while abstracting from informational noise in peer-ability signals, schooling-market search frictions or matching dynamics, on-the-job and familial human-capital externalities, quantitative calibration of parameters, and any normative or policy counterfactuals, each of which could attenuate but is unlikely to overturn the core peer-sorting motive, and so we leave them for future extensions.

Anticipating that high-ability peers accelerate learning while lower-ability peers impose a drag, rational households tend to prioritize the school environment that maximizes their child's expected skill trajectory. High-efficiency families are therefore encouraged to cluster together to capture positive peer gains, whereas low-efficiency families have little incentive to forgo integration as they benefit from mixing with stronger learners. This asymmetry in the desire to segregate—the strong sorting motive of high-ability households versus the weaker motive of low-ability ones—echoes Lucas & Moll's (2014) insights on optimal learning environments and complements the early-life skill formation literature (Cunha, Heckman, & Schennach, 2010; Heckman & Masterov, 2007; Neal, 2018).

The paper unfolds as follows. In Section 2, we introduce a discrete-time overlapping-generations framework in which human-capital accumulation is driven by the prospect of future income; learning depends on previously acquired knowledge and the time each child devotes to study. In Section 3, we reformulate this setup in continuous time, construct the representative household's Hamiltonian, and derive the existence and uniqueness conditions for the isolated learner's optimal control problem. Section 4 then examines the stability of that optimum, while Section 5 fully characterizes the optimal time-allocation path for both the decentralized household and the social planner.

Building on this benchmark, Section 6 embeds a positive peer externality into the learning function, solves for both equilibrium and planner allocations, and shows how they diverge once classmates' abilities matter. To trace those dynamics back to the no-externality case, Section 7 applies Oniki's and Caputo's comparative-dynamics methods. Finally, Section 8 reflects on our core insights, aligns them with empirical patterns of school clustering, and flags avenues for future work.

2. An overlapping generations model

In a family, each generation lives for two periods, so after being young it gets old. In any time t, there are two generations, young and old, living simultaneously. In the first period of her life, the individual learns at school to accumulate human capital that she uses in the second part of her life when working. As a capital transfer from the previous generation, the member of the young generation living at time t (called as the member of generation t when young) enters the school system with a knowledge level h_t^{α} with $0 < \alpha < 1$. The human capital of the old generation living at time t (or the member of generation $t-1$ when old) is h_t, so the parameter α determines the

fraction of human capital transferred from the old to the young that the young generation uses to learn new skills. It must hold that $h_t^{\alpha} < h_t$ for all t as the young generation cannot possess higher capital level when entering the school system than their parents do. With $0 < \alpha < 1$ it implies that $h_0 > 1$, so there is a barrier for a family to enter school the first time in terms of a minimum required knowledge. This assumption captures the fact that children are supposed to have some skills and knowledge before starting formal education at the required age. This lower bound $h_0 > 1$ on human capital will be reinforced further in the subsequent sections. Depending on this capital transfer, a member of generation t when young realizes utility according to a strictly increasing and strictly concave school performance function. For simplicity, this function is assumed to be $\ln(h_t^{\alpha}) = \alpha \ln(h_t)$. Human capital accumulation is described with a learning function $h_{t+1} = g(h_t, u_t)$ according to which learning requires knowledge, h_t, and time, u_t. For the young individual living in t, time investment is the decision variable of which $u_t \in [0,1]$ must hold. Time investment is foregone leisure, so u_t is not included in the time someone spends learning at school. Accordingly, the case $u_t = 0$ means that time investment in learning is minimal and constrained to school activities. Function g is time-invariant, so the same learning capabilities characterize the whole dynasty. Time investment is fraught with costs described with a strictly increasing and strictly convex cost function. For simplicity, a quadratic function $-Cu_t^2$ is assumed to describe the disutility of learning with cost parameter $C > 0$. Obviously, the higher C is, the higher level of disutility (or, equivalently, a lower level of utility) is realized from the same ratio of time.

Learning has a twofold purpose. Since in the second period of her life the individual works, her decision on time investment u_t depends partly on the present value of the expected return on human capital. This expected value is described with a strictly increasing and strictly concave function. For simplicity, this function is assumed to take the form $\beta B \ln(h_{t+1})$, in which $\beta > 0$ is the one-period discount factor, and $B > 0$ is a parameter proportional to the value of the income stream generated by then human capital h_{t+1} accumulated for the second part of her life. A young individual also considers the expected utility of her later child, given as $\beta\, U_{t+1}^*(h_{t+1})$, which is the time-$t$ value of the maximum attainable utility from a given level of human capital owned in the second period of life from which she of course transfers a fraction h_{t+1}^{α}. In other words, the young individual living in t deciding on u_t hence h_{t+1} also takes into account how much utility the next generation can realize from any given h_{t+1} as her utility directly depends on her child's utility. Therefore, assuming $h_1 = g(h_0, u_0)$, the problem of the decision maker in time 0 is

$$U_0^*(h_0) = \max_{u_0 \in [0,1]} \{\alpha \ln(h_0) - Cu_0^2 + \beta B \ln(h_1) + \beta U_1^*(h_1)\} \qquad (1)$$

as the member of generation 1 when young makes her investment decision u_1 in the next period, $t = 1$, when living simultaneously with the induvial having made her decision in time 0. Capital transfer h_0^{α} is wholly predetermined in terms of the problem of the time-zero decision maker and comes from the currently living old member of generation -1.

The optimum levels of u_0 and h_1, given h_0, also depends on $U_1^*(h_1)$, so the member of generation 0 when young must also consider the problem of what utility her optimizing child will attain from any feasible h_1. Therefore, her problem widens and takes the form

$$U_0^*(h_0) = \max_{u_0\in[0,1],u_1\in[0,1]}\{\alpha\ln(h_0) - Cu_0^2 + \beta B\ln(h_1) + \beta[\alpha\ln(h_1) - Cu_1^2 + \beta B\ln(h_2) + \beta U_2^*(h_2)]\} =$$
$$\max_{u_0\in[0,1],u_1\in[0,1]}\{\alpha\ln(h_0) - Cu_0^2 + \beta B\ln(h_1) + \beta\alpha\ln(h_1) - \beta Cu_1^2 + \beta^2 B\ln(h_2) + \beta^2 U_2^*(h_2)\}.$$

subject to $h_{t+1} = g(h_t, u_t)$. Going on with this pattern, eventually, it turns out that, to solve its investment problem, the young individual living in time 0 must find the optimal investment and capital paths for the whole family as her utility depends on the utility realized by all the subsequent generations. Her problem can be written in the sequential form

$$U_0^*(h_0) = \max_{\{u_t\}_{t=0}^{\infty}|u_t\in[0,1]}\{\alpha\ln(h_0) - Cu_0^2 + \beta B\ln(h_1) + \beta\alpha\ln(h_1) - \beta Cu_1^2 + \beta^2 B\ln(h_2) +$$
$$\beta^2\alpha\ln(h_2) - \beta^2 Cu_2^2 + \beta^3 B\ln(h_3) + \beta^3\alpha\ln(h_3) - \beta^3 Cu_3^2 + \beta^4 B\ln(h_4) + \cdots\}.$$

The agent has control of u_0 hence h_1 only, so all the subsequent generations make their own decisions. Her child thus needs to find the optimal u_1 hence h_2 with given h_1, but this decision is the same as she expected her child to make out of a given h_1 that she determined optimally for herself. When deciding on u_0 hence h_1 she thus foresaw her child's decision on u_1 hence h_2, and so on with all the future generations, so the whole dynasty makes the same decisions as the member of generation 0 when young. We can thus solve the overlapping generations problem as the problem of an infinitely living representative agent, in line with Barro's (1974) prototypical application.

The problem and its optimal solution, if it exists, do not change when a constant term is added to the objective function, or when the objective function is scaled with a positive constant.[1] Accordingly, we can add $B\ln(h_0)$ and then multiply the augmented objective function with $1/(B+\alpha) > 0$ that yields

$$U_0^*(h_0) = \max_{\{u_t\}_{t=0}^{\infty}|u_t\in[0,1]}\left\{\ln(h_0) - \frac{C}{B+\alpha}u_0^2 + \beta\ln(h_1) - \beta\frac{C}{B+\alpha}u_1^2 \right.$$
$$\left. + \beta^2\ln(h_2) - \beta^2\frac{C}{B+\alpha}u_2^2 + \beta^3\ln(h_3) - \beta^3\frac{C}{B+\alpha}u_3^2 + \cdots\right\}$$

that is equivalent to

$$\max_{\{u_t\}_{t=0}^{\infty}}\sum_{t=0}^{\infty}\beta^t\left[\ln(h_t) - \frac{c}{2}u_t^2\right] \tag{2}$$

[1] Bertsekas (1976, p. 5) lays down the principle that an affine transformation of a payoff function that preserves the orderings of alternative policies yields the same solution. Likewise, if an optimal solution is regarded as an element of a sequence space such as $\{u_t\}_{t=0}^{\infty} \in l_\infty$, then the return function G is a functional on this space. If there is some program u^* such that $G(u) < G(u^*)$ where $G(u^*) = \sup\{G(x): u \in l_\infty\}$ for any feasible program u, then with $a, b \in \mathbb{R}$ and $a > 0$ it holds that $aG(u) + b < aG(u^*) + b$ that can also be written as $\tilde{G}(u) < \tilde{G}(u^*)$, so the same solution emerges before and after the linear transformation of the payoff function.

subject to $h_{t+1} = g(h_t, u_t)$ in which

$$\frac{C}{B+\alpha} \equiv \frac{c}{2} > 0$$

for computational reasons becoming obvious later. In what follows, we solve the continuous-time version of (2) with different specifications of the learning function g. The dynamic problem embedded in (2) is to find an optimal path for the capital stock the benefits of which can only be realized only at the price of some disutility. Tomorrow's returns can only be increased by sacrificing some returns today.

3. An infinitely living representative agent in continuous time

The member of generation 0 when young is assumed to use human capital h_0^{α} with diminishing marginal product and time investment $u_0 \in [0,1]$ in learning. Human capital h_{t+1} consists of two parts. The first part comes from learning with time investment and the use of the knowledge transfer, and arises as a stock containing only new skills not inherited from the previous generation. This part is $A(h_t^{\alpha})^{\zeta} u_t$ in which $0 < \zeta < 1$ describes the diminishing marginal productivity of human capital in learning, and $A > 0$. The other part comes from inherited knowledge subject to the depreciation of human capital from one generation to the next. If $\delta > 0$ is the depreciation rate determining how skills and knowledge lose their relevance from one generation to the next, then this part is simply $(1-\delta)h_t$. This component is acquired during school hours and requires no additional private time investment. Therefore, the learning function g takes the

$$h_{t+1} = A(h_t^{\alpha})^{\zeta} u_t + (1-\delta)h_t$$

form that implies

$$h_{t+1} - h_t = A(h_t^{\alpha})^{\zeta} u_t - \delta h_t \qquad (3)$$

which is a discrete-time version of Lucas's (1988, pp. 19, Eq. (13)) learning function that is the same linear in time investment, but exhibiting diminishing marginal product of human capital. When $u_t = 0$, $h_{t+1} = (1-\delta)h_t$, so human capital in the next period is equal to the relevant part of the stock of knowledge remaining from the current period with no new knowledge added. The learning function in (3) remains compatible with Lucas's (1988) idea of continuous technological progress even if (3) might imply (and, as we shall see it soon, it actually does imply) a stable optimum stock of knowledge. At this stable level, a fraction of human capital disappears due to depreciation and is replaced with new and previously not possessed knowledge. Therefore, knowledge and technology can continuously improve even though the stock of knowledge of each subsequent generation is approaching toward a constant level.[2]

[2] As Lucas pointed out in the drafts he made for his 1988 endogenous growth theory paper (Lucas papers, Box 15), knowledge is always someone's knowledge hence it is rightly regarded as personal. Since the capacity of the human mind

Diminishing marginal product of knowledge stems from the restriction on ζ. A is responsible for the productivity of time and knowledge taken as the factors of production in knowledge creation and describes what growth rate is achievable from unit capital and unit time investment. Note that (3) can be rewritten by using the fact that $(h_t^{\alpha})^{\zeta} \equiv h_t^{\eta}$ in which $0 < \eta < 1$ applies by assumption following from our previous restrictions on α and ζ. When learning, the time t decision maker can thus be taken as to consider how much capital to accumulate for the next period out of h_t and not h_t^{α}, so at the end of the day the learning function in (3) turns out to take the form

$$h_{t+1} - h_t = A h_t^{\eta} u_t - \delta h_t \tag{4}$$

as, by definition, the sequence $\{h_t\}_{t=0}^{\infty}$ of the human capital of the family keeps track of the stock of knowledge owned by the old member of each generation. In continuous time, the objective function in (2) becomes

$$\max_{u(t)} \int_0^{\infty} e^{-\rho t} \left[\ln h(t) - \frac{c}{2} u^2(t) \right] dt \tag{5}$$

subject to

$$\dot{h}(t) = A\, h^{\eta}(t)\, u(t) - \delta h(t) \tag{6}$$

as the continuous time version of (4) where $h(t)$ and $u(t)$ denote the time paths of human capital and time investment with the restriction $0 \leq u(t) \leq 1$. Even if higher learning abilities run in a family that may put it on a more aggressively increasing capital profile, the high level of the subjective disutility might prevent the dynasty from fulfilling its learning potential. Note that there is a trade-off between α and ζ, so a family can offset the effects of a lower α on human capital accumulation with a higher ζ. If α is assumed to be common to all families, then differences in η solely come from differences in productivity, measured with differences in ζ. Therefore, in the subsequent analysis we do not need the distinction between α, ζ, and η, so all attention is paid to differences in η only. Parameter η is construed to include all genetical and institutional factors (such as cozy study spaces at home, supportive family environment, etc.) affecting how productively an individual uses previously acquired knowledge in human capital accumulation, especially how the stock of human capital is exposed to diminishing marginal product. The parameter is thus meant to incorporate some room for compensation between hereditary and

---

is limited, it cannot contain, process and apply an infinite amount of knowledge. Practically, the idea of the learning function in (3) is that a certain fraction of current knowledge gets out of use. Even if it remains retrievable from, for instance, documentation or other forms of technical descriptions, nobody owns it as it proved to be impractical for the current generation. The classic example of the idea is the Saturn vehicle of the Apollo Program of NASA the manufacturing of which was based on technologies and supply chains that had disappeared by today (Bilstein, 1980). The manufacturing procedures are still available in the documentation, but nobody knows today *by heart* how to build a Saturn.

institutional factors (Heckman, 2009). A family can set off the flaws of disadvantageous genetics with investment in the institutional factors that increase η; or, in other words, the positive effects of favorable genetical factors can be compromised and undermined by the impact of an unsupportive family environment.

When we subsequently form a class of students to make learning a social process, the positive effect of the class environment in which the individual is no longer learning alone is modelled by an increase in η in the fashion of Lucas (1988). Parameter η by itself is unexperienceable to the social environment of a family, while in a social setting it plays a crucial role in the selection of desired classmates to children. As genetics remain imperceptible to the common eye, rational families are construed to infer η from the socio-economic status of a family. In reality, inferring η from signals like socioeconomic status is fraught with noises that must be considered in any model that is supposed to give account of peer selection and search leading to segregation. The present model, however, explains the *drive to segregate* not segregation, so noises are assumed away.

Together, parameters A and η describe learning abilities of the individual and can be identified as indicators how productive the individual is in learning, and how fast she can accumulate human capital. The model thus draws a sharp distinction between the level of knowledge (how smart a student is) and her learning abilities (how fast and productively she can learn). On this ground, the maximum attainable level of human capital

$$\bar{h} = \left(\frac{A}{\delta}\right)^{\frac{1}{1-\eta}} \tag{7}$$

emerges at the intersection of the curve of maximum production of new knowledge attained at $u = 1$ at any $h(t)$, and capital loss due to depreciation. Due to diminishing marginal product of human capital in its own production such a $\bar{h}$ exists and it is unique. Obviously, $\bar{h}$ is determined by the three parameters of the learning function.

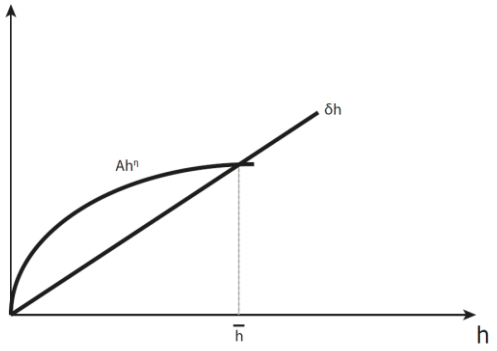


Fig. 1: Determination of the maximum capital stock. The curve Ah^{η} comes from Eq. (6) with the maximum time investment where $u = 1$ applies.

A limitation of the model must be highlighted at this early phase. Given (6), the rate of increase of h is dependent on η, and we expect $\dot{h}'_\eta > 0$ as, at any given level of h, depending on time investment, the student reaches a certain rate of capital accumulation. By requiring $\dot{h}'_\eta > 0$ we expect a student having higher learning abilities to reach a faster accumulation rate *ceteris paribus*. However, by direct calculation we have

$$\dot{h}'_\eta = Ah^\eta u \ln(h),$$

which is negative if $0 < h < 1$, while $\dot{h}'_\eta > 0$ for any $h > 1$. In the former case, maintaining a certain rate of human capital accumulation at the same knowledge level would require a student having higher η to invest *more* time than her classmate with a lower η. Consequently, the model does not seem to adequately describe human capital accumulation at very low levels of knowledge. It causes no serious problem, however, as the model is intended to deal with school education to which students are exposed after accumulating (or required to have accumulated) a minimum required level of h. Additionally, in empirical work it does no harm to scale the initial levels of knowledge to require $h_0 > 1$. The process of pre-school learning might be addressed with another framework.

The current-value Hamiltonian of the problem is

$$H(h;\theta;u) = \ln h(t) - \frac{c}{2}u^2(t) + \theta(t)[\,Ah^\eta(t)u(t) - \delta h(t)] \tag{8}$$

where θ denotes the costate variable of h. Getting the necessary conditions is standard and yields the three-equation system

$$u = \frac{A}{c}\theta h^\eta \tag{9}$$

$$\dot{\theta} = \theta(\rho + \delta - A\eta h^{\eta-1}u) - \frac{1}{h} \tag{10}$$

$$\dot{h} = Ah^\eta u - \delta h \tag{11}$$

As h cannot leave the state space $[0;\bar{h}]$, there must be a steady-state capital stock

$$h^* = \left[\frac{\frac{A^2}{\delta c}}{\rho + \delta(1-\eta)}\right]^{\frac{1}{2-2\eta}}$$

and costate

$$\theta^* = \frac{\delta c}{A^2}(h^*)^{1-2\eta}$$

that together also determine a steady-state time investment u^* through (9). (Appendix A)

4. Stability of the optimal level of knowledge

Stability of the optimal knowledge requires h^* to be maintainable, so $h^* \leq \bar{h}$ must apply as a necessary condition of stability. As it turns out, both $h^* \leq \bar{h}$ and $u^* \leq 1$ follow if

$$\frac{\delta}{\rho + \delta(1-\eta)} \leq c \tag{12}$$

holds as a condition. The requirement in (12) is straightforward to interpret as a condition according to which the agent does not regard time investment so cheap and undemanding that she would want to invest $u > 1$ which is impossible. Just like in Ben-Porath's (1967) human capital accumulation model, such a case can practically be avoided by activating the upper bound $u \leq 1$ for time investment whenever $u > 1$ proves to be optimal. As we shall see in the following section on the saddle path analysis, $\dot{u} < 0$ can commonly be expected all over the path towards h^*, and even if a case where $u > 1$ cannot be excluded, it occurs only at the early phases of learning with low level and high value of capital (the latter is expressed by the path of the costate variable), and u eventually reduces into the $u \leq 1$ range. The $u \not> 1$ threshold may therefore be active only temporary.

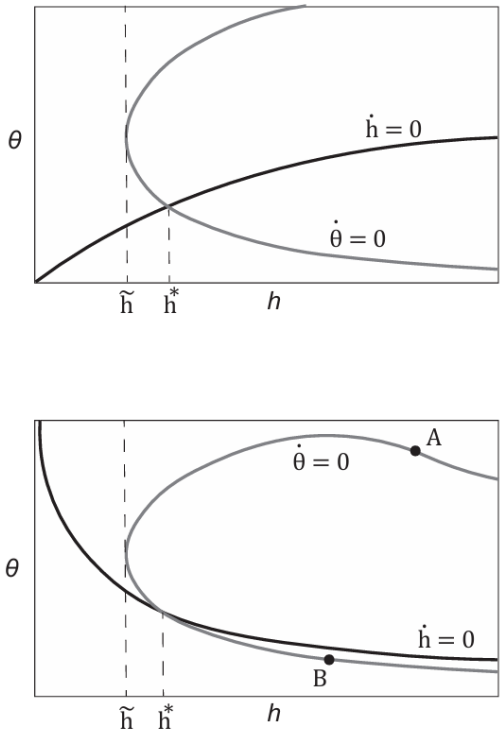


Fig. 2: Typical shapes of the nullclines with $0 < \eta < 0.5$ on the upper pane, and $0.5 < \eta < 1$ on the lower pane. The lower pane also includes a schematic representation of the ambiguity of the $0.5 < \eta < 1$ case. In point *A*, moving to the left we eventually get to the right of the $\dot{\theta} = 0$ nullcline. Likewise, in point *B*, moving upwards we eventually get under the same nullcline.

The $\dot{h} = 0$ and $\dot{\theta} = 0$ nullclines (see Appendix A and B, respectively) take different shapes depending on η. Due to stability reasons, the $0.5 < \eta < 1$ case proves to be irrelevant. However, for $0 < \eta < 0.5$ the system is saddle-path stable, so to any $1 < h_0 < h^*$ there is a unique $\theta(0)$ from which the stock of knowledge converges toward the steady-state solution h^* (Appendix C). A simple calibration method is available to handle the limitation in the assumption $0.5 < \eta < 1$. Supposing h_0 (as the starting value) and h_1 (as the terminal value) are estimated by school test data for a class or cohort, and time investment $u_0 \in [0,1]$, as a summary figure of estimated time investment during the learning period, is also known from survey results, we have two free parameters, A and η, to calibrate. Through (4) it is straightforward to find individual η parameters for a given $A > 0$ that is assumed to be common for all students in the group. From (4), it is obvious that η inversely depends on A, so in the simplest case, starting with a sufficiently low $A > 0$, with increasing A step by step, it is possible to iteratively find the lowest A that squeezes the highest η under 0.5 while keeping the $0 < \eta$ assumption. In more complicated cases, however, we can squeeze the highest η under 0.5 only at the price of violating the $0 < \eta$ assumption, so some individual η parameters turn out to be negative. This case shows that we might have such a diverse cohort in which the assumption of a single common A cannot be maintained. This is exactly the difference in As that we assume away for simplicity.

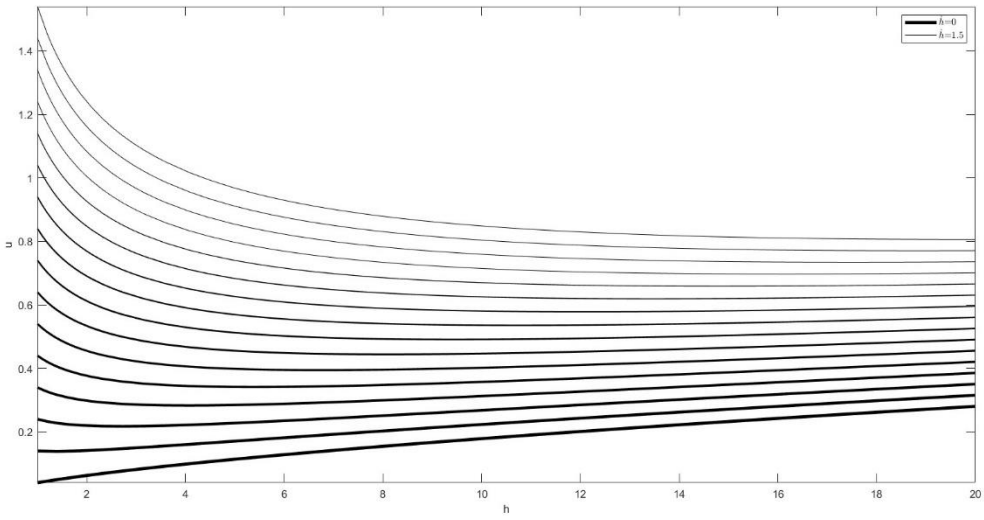
5. Saddle path analysis

As learning is a timely process, utility of the representative agent depends on both the level of the steady-state capital stock (h^*), its maintenance costs (u^*) on the one hand, and the optimum path towards the equilibrium point on the other hand. Therefore, attention must also be paid to the characteristics of this path. In this section we characterize the optimum path of the capital stock as closely as possible. As the model is built on a convex cost function, a most important question regards the strict concavity of the optimum path. There exists a huge literature, initiated by the seminal work of Eisner and Strotz (1962) and its immediate sequel by the early Lucas (1965a; 1965b; 1965c), on how convex costs encourage the agent to postpone investment. As they argue, at every point in time the agent might choose a sharp increase in the state variable, but it always proves to be cheaper to pick a slower pace of expanding, hence concavity of the optimal capital path emerges. In the present model, however, the sign of $\ddot{h}$ is not clearly negative, so some convexity in the early phases of human capital accumulation cannot be excluded. Such a convex section signals that the representative agent is interested in quick learning.

We can realize that the model establishes a close relationship between η and $\theta(0)$. Given h_0, a higher η ensures access to a higher h^* and increases the value $\theta(0)$ (Appendix D). To see this, it is useful to realize that the isoquant curves of the learning function in (6) shift at different values of h in different ways. From (6), the dynamic isoquants take the

$$\frac{\dot{h} - \delta h}{A h^{\eta}} \equiv f(h) = u$$

form for some $\dot{h} > 0$, and for a starting value h_0 that is sufficiently close to 1 an increase in η does not make it much easier for the agent to reach and maintain a given level of $\dot{h}$. While a higher learning parameter η elevates h^*, starting from a sufficiently low $h_0 > 1$, the agent cannot rely on her improved learning abilities in achieving higher levels of $\dot{h}$ in the early stage of human capital accumulation. As a change in η affects the curvatures of the isoquant curves in this region of h (see Fig. 3 below), in order to achieve the higher h^* with higher accumulation rates $\dot{h}$ the agent needs to invest more time. Therefore, an increase in η does not ease the scarcity of h_0, so higher values of η tend to increase $\theta(0)$, and we have $\theta(0)_{\eta} > 0$ at low levels of h_0. By contrast, an increase in A shifts a whole isoquant *downwards* making for the agent easier to maintain a given level of $\dot{h}$. Therefore, even if h^* increases and h_0 becomes scarcer with A, an increase in A also eases this scarcity of human capital at low levels of h_0, so $\theta(0)_A < 0$. In later phases, an increase in η proves to be more useful as compared to an increase in A by making it possible for the agent to maintain given levels of $\dot{h}$ at smaller time investments. Simulation results endorse this pattern.



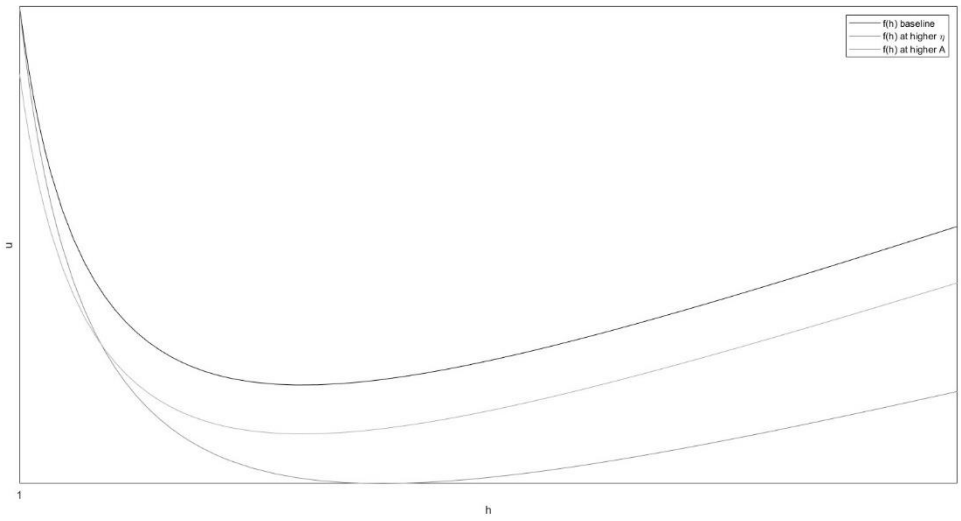


Fig. 3: The isoquant map of the learning function (upper pane, where $A = 1$ and $\delta = 0.04$); and the typical effects of a change in η or A on the shape of the learning function (lower pane, parameters being the same). Most importantly, as the lower pane shows, at low levels of $h_0 > 1$ an increase in A reduces the time needed for maintaining a given value of $\dot{h}$ more than an increase in η.

These insights imply that by decreasing $h_0 > 1$ and increasing η we may be able to force the representative agent to more aggressively learn hence to find a starting value for the optimum human capital accumulation process that has an initial convex section. Fig. 4 shows that increasing η in fact causes initial convexity which fades away at lower values of η.

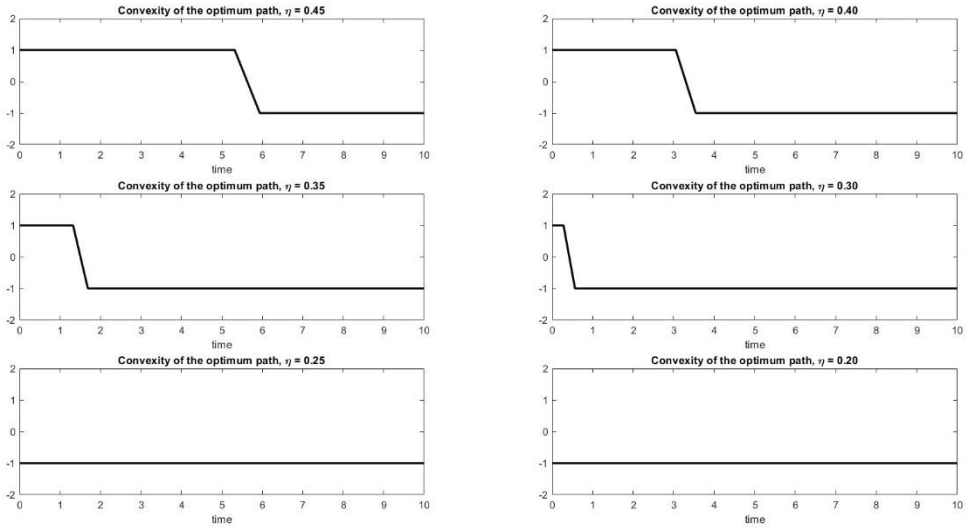


Fig. 4: Convex sections of the optimum path of h at different values of η. The vector of further parameters are $[A, c, \delta, \rho] = [1, 5, 0.04, 0.05]$. Appendix D shows the formal criterion of convexity whose sign we can use in simulation. Positive sign, indicated by the signum function applied to the criterion, means convexity.

It is easy to show that from (9)

$$\dot{u} = \frac{A}{c}\left[\dot{\theta} h^{\eta} + \eta\theta h^{\eta-1}\dot{h}\right] \tag{13}$$

follows which yields

$$\dot{u} = \frac{A[\rho + \delta(1-\eta)]}{c}\theta h^{\eta} - \frac{A}{c}h^{\eta-1} \tag{14}$$

which is negative hence $u(t)$ is decreasing in time as long as

$$\theta < \frac{1}{\rho+\delta(1-\eta)}\frac{1}{h}. \tag{15}$$

Time investment thus depends on the relationship between θ and h, and $u(t)$ might increase in time if the inequality gets reversed in (15). As the first term on the right is expected to be relatively high, in all these cases $\dot{u} < 0$. Furthermore, by direct calculation it is straightforward to point out that maintaining the optimum capital stock requires more time to be invested by students having higher learning abilities, albeit this disutility is more than offset by the higher values of h^* accompanying higher values of η.

Note that an initial convex capital path might involve sections where $u(t) > 1$ occurs, so in such cases the upper bound $u(t) \leq 1$ must be active. In such cases, uniqueness and saddle path stability still apply (Appendix E).

6. Learning as a social activity

After characterizing the problem of the infinitely living representative agent taken as an isolated family, in Sections 6–7 we turn human capital accumulation into a social problem. Practically, we do this by introducing an external effect into the learning function to describe how learning abilities interplay in communities. With an external effect, however, we must distinguish equilibrium and optimum solutions. Following Lucas (1988) and Romer (1986), the equilibrium paths of $h(t)$ and $u(t)$ directly result from the solutions of the isolated problems. In the equilibrium, families forming a community optimize individually. Even if the individual representing a dynasty knows that her own learning process is affected by how others accumulate human capital, having no control over these processes, she regards these paths as exogenously given. Note that there is no market interaction in this setting: Human capital is not traded, it only externally affects the learning processes in a community. The optimum allocation, by contrast, is established by the social planner, now taken as a schoolteacher having full control over the time investment of each student in the class, who does not need to take the individual capital paths as exogenously given but simultaneously plans optimal paths for all the members in the community.

The external effect modifies the learning function of the individual in a way that bears close resemblance to the way Lucas (1988, p. 18) introduced the external effect of knowledge into the

production of commodities. Instead of describing average knowledge with an algebraic average, however, we use the geometric mean to preserve the multiplicative relationship in the learning function. To keep things simple, we work with two individuals denoted by 1 and 2, so average knowledge level is

$$h_a(t) = \sqrt{h_1(t)\, h_2(t)}$$

and the production function in (6) becomes

$$\begin{aligned} \dot{h}_1 &= A h_a^{\gamma} h_1^{\eta_1} u_1 - \delta h_1, \\ \dot{h}_2 &= A h_a^{\gamma} h_2^{\eta_2} u_2 - \delta h_2 \end{aligned} \tag{16}$$

in which we allow η_1 and η_1 to be different, and $\gamma > 0$ is the parameter that captures the external effect of human capital accumulation. Obviously, γ is straightforward to interpret as reflecting the overall quality of the learning environment. Simply put, a better school is signaled with a higher γ. Further specifications will be introduced later. When $\gamma = 0$, we have the problem of the isolated individual. Without losing generality, we may assume that $0.5 > \eta_1 > \eta_2 > 0$, so dynasty 1 has higher learning abilities than the other family in the social setting. Note that $h_a(t)$ depends on the personal knowledge level for each individual, so it is straightforward to rewrite (16) as

$$\begin{aligned} \dot{h}_1 &= A h_1^{\frac{\gamma}{2}} h_2^{\frac{\gamma}{2}} h_1^{\eta_1} u_1 - \delta h_1 = A h_2^{\frac{\gamma}{2}} h_1^{\eta_1 + \frac{\gamma}{2}} u_1 - \delta h_1, \\ \dot{h}_2 &= A h_1^{\frac{\gamma}{2}} h_2^{\frac{\gamma}{2}} h_2^{\eta_2} u_2 - \delta h_2 = A h_1^{\frac{\gamma}{2}} h_2^{\eta_2 + \frac{\gamma}{2}} u_2 - \delta h_2 \end{aligned} \tag{17}$$

in which capital stock and learning parameter of the other student is exogenously given to each individual. The problem can thus be reduced to the case of the isolated individual—the exogenously given part $h_i^{\frac{\gamma}{2}}$ sticks to the exogenously given parameter A, and the external effect elevates the learning parameter of each individual to $0 < \eta_j + \frac{\gamma}{2} < 0.5$, where $i = 1,2$ and $j = 1,2$ and $i \neq j$. Equilibrium knowledge levels are finite as there are upper bounds on the set of feasible capital stocks (Appendix F).

Under the previous assumptions, the Hamiltonians of the individuals in the equilibrium setting are

$$\widehat{H}_1(h_1, \theta_1, u_1) = \ln(h_1) - \frac{c}{2}(u_1)^2 + \theta_1 \left(A h_2^{\frac{\gamma}{2}} h_1^{\eta_1 + \frac{\gamma}{2}} u_1 - \delta h_1 \right), \tag{18a}$$

$$\widehat{H}_2(h_2, \theta_2, u_2) = \ln(h_2) - \frac{c}{2}(u_2)^2 + \theta_2 \left(A h_1^{\frac{\gamma}{2}} h_2^{\eta_2 + \frac{\gamma}{2}} u_2 - \delta h_2 \right). \tag{18b}$$

It is easy to realize that each problem is a version of (8) with new exogenous parameters $A h_1^{\frac{\gamma}{2}}(t)$, $A h_2^{\frac{\gamma}{2}}(t)$, $\eta_1 + \frac{\gamma}{2}$ and $\eta_2 + \frac{\gamma}{2}$ in the learning functions. We can thus use the fact that each individual

regards the capital stock of the fellow student as exogenously given. Consequently, the equilibrium stocks h_1^e and h_2^e are

$$h_1^e = \left[\frac{\frac{A^2}{\delta c}(h_2^e)^\gamma}{\rho + \delta\left(1 - \eta_1 - \frac{\gamma}{2}\right)} \right]^{\frac{1}{2-2\left(\eta_1+\frac{\gamma}{2}\right)}} \tag{19a}$$

$$h_2^e = \left[\frac{\frac{A^2}{\delta c}(h_1^e)^\gamma}{\rho + \delta\left(1 - \eta_2 - \frac{\gamma}{2}\right)} \right]^{\frac{1}{2-2\left(\eta_2+\frac{\gamma}{2}\right)}} \tag{19b}$$

and consequently

$$\theta_1^e = \frac{\delta c}{A^2(h_2^e)^\gamma} \left[\frac{\frac{A^2}{\delta c}(h_2^e)^\gamma}{\rho + \delta\left(1 - \eta_1 - \frac{\gamma}{2}\right)} \right]^{\frac{1-2\left(\eta_1+\frac{\gamma}{2}\right)}{2-2\left(\eta_1+\frac{\gamma}{2}\right)}}$$

and symmetrically

$$\theta_2^e = \frac{\delta c}{A^2(h_1^e)^\gamma} \left[\frac{\frac{A^2}{\delta c}(h_1^e)^\gamma}{\rho + \delta\left(1 - \eta_2 - \frac{\gamma}{2}\right)} \right]^{\frac{1-2\left(\eta_2+\frac{\gamma}{2}\right)}{2-2\left(\eta_2+\frac{\gamma}{2}\right)}}$$

in which, therefore, the equilibrium knowledge levels thus depend on each other. Appendix A helps the reader to realize the similarity with the solution to the isolated case. To find the equilibrium stocks, we have a fixed-point problem, hence we can follow a similar strategy that we applied in the case of the maximum levels $\bar{h}_1$ and $\bar{h}_2$. After manipulation, we have

$$h_1^e = \left[\frac{\frac{A^2}{\delta c}}{\rho + \delta\left(1 - \eta_1 - \frac{\gamma}{2}\right)} \right]^{\frac{2-2\eta_2-\gamma}{(2\eta_1+\gamma-2)(2\eta_2+\gamma-2)-\gamma^2}} \left[\frac{\frac{A^2}{\delta c}}{\rho + \delta\left(1 - \eta_2 - \frac{\gamma}{2}\right)} \right]^{\frac{\gamma}{(2\eta_1+\gamma-2)(2\eta_2+\gamma-2)-\gamma^2}}, \tag{20a}$$

$$h_2^e = \left[\frac{\frac{A^2}{\delta c}}{\rho + \delta\left(1 - \eta_2 - \frac{\gamma}{2}\right)} \right]^{\frac{2-2\eta_1-\gamma}{(2\eta_1+\gamma-2)(2\eta_2+\gamma-2)-\gamma^2}} \left[\frac{\frac{A^2}{\delta c}}{\rho + \delta\left(1 - \eta_1 - \frac{\gamma}{2}\right)} \right]^{\frac{\gamma}{(2\eta_1+\gamma-2)(2\eta_2+\gamma-2)-\gamma^2}}, \tag{20b}$$

as the closed-form equilibrium knowledge levels according to which h_1^e depends on η_2, so the student having inferior learning abilities affects the equilibrium knowledge level of her fellow classmate. Obviously, when $\eta_1 = \eta_2 \equiv \eta$, $h_1^e = h_2^e$. A key question regards the direction of this effect. It can be shown that both equilibrium capital stocks h_1^e and h_2^e positively depend on the learning parameters η_1 and η_2 under a sufficiently wide range of parameter configurations (Appendix G).

One can realize that the equilibrium problem in the class can be described as an infinite chain of isolated problems. From the viewpoint of Student 1, the classmate's capital stock sticks to A as $A\big(h_2(t)\big)^{\frac{\gamma}{2}}$ in the learning function, the equilibrium problem of Student 1 is thus analogous to an isolated problem with an exogenously increasing A. Remember that at any appropriate parameter configuration, the isolated case is saddle-path stable with $\dot{h} > 0$, so with any exogenous $h_2(0)$, knowledge level $h_1(0)$ starts growing hence $\dot{h}_1(0) > 0$ and vice versa. At the start of the learning process where $t = 0$, with $h_1(0)$ and $h_2(0)$ given, Student 1 starts converging along a stable manifold to

$$\vec{h}_1(0) = \left[\frac{\left(A\left(h_2^{\frac{\gamma}{2}}(0) \right) \right)^2}{\frac{\delta c}{\rho + \delta\left(1 - \eta_1 - \frac{\gamma}{2}\right)}} \right]^{\frac{1}{2-2\left(\eta_1+\frac{\gamma}{2}\right)}}$$

as a temporary capital endpoint or center of gravity that is a function of $h_2(0)$ as the analogue of (19a). The temporary capital endpoint for Student 2 is similar, with $h_1^{\frac{\gamma}{2}}(0)$ involved. Given that h_2 is also increasing by the same saddle-path stability, we have temporary endpoints of convergence as $\vec{h}_1(t)$ and $\vec{\theta}_1(t)$, and $\vec{h}_2(t)$ and $\vec{\theta}_2(t)$ at any time t. As we describe the equilibrium case as the analogue of the optimum in the isolated case, it is straightforward to point out that $\vec{h}_i(t)$ is increasing while $\vec{\theta}_i(t)$ is decreasing in time—suffice it to differentiate h^* and θ^* with respect to A in the isolated case, and $Ah_i^{\frac{\gamma}{2}}$ is certainly growing for both agents for $\dot{h}_i > 0$. Any pair $\vec{h}_i(t)$ and $\vec{\theta}_i(t)$ at any t establish a temporary steady state point to which there always belongs a unique stable manifold. So, students are converging to moving targets, tending toward the lower-right corner of the phase diagrams in Fig. 10 in Appendix C, with $\lim_{t\to\infty} \vec{h}_i(t) = h_i^e$ as described in (19a) and (19b). The phase diagram of the equilibrium problem of the individual is a combination of the phase diagrams depicting the problem at any t, in which we converge in the state toward h_i^e while shifting from one temporary saddle path to the next. It is also easy to show that h_1^e and h_2^e are the same as h^* in the isolated case when solved with an A increased by $(h_i^e)^{\frac{\gamma}{2}}$, and a productivity parameter η increased by $\gamma/2$. Note that in this section we only establish the claim that at any point in time for each agent there exists a stable manifold, and the agent converges towards h_i^e by continually shifting from one saddle path to another, but disregard how this convergence affects the optimum paths of the capital stocks. We address this aspect of the problem in Section 7 in a comparative dynamics context.

As we reduce the equilibrium dynamics to the case of the isolated individual, it is useful to examine how the endpoints of convergence shift. Regarding the differentials of h^* and θ^*, note that $h^*_\eta > 0$ holds as long as

$$\frac{\frac{A^2}{\delta c}}{\rho + \delta(1-\eta)} > 1$$

which is the only relevant case in order for $h^* > 1$ to emerge, so higher learning abilities embedded in higher values for η lead to higher optimum knowledge level. For normal parameter values $\theta^*_\eta < 0$ also holds that expresses the intuition that the higher learning abilities an individual has, the less precious she regards an additional unit of knowledge. To see this, we start by rewriting the expression for θ^* as

$$\theta^* = [\rho + \delta(1-\eta)]^{\frac{2\eta-1}{2-2\eta}} \left(\frac{\delta c}{A^2}\right)^{\frac{1}{2-2\eta}}.$$

A sufficient but not necessary condition for θ^* to be a decreasing function of η is when each term on the right is a decreasing function of η. The sign of the differential of the first term is negative if

$$\ln[\rho + \delta(1-\eta)] + \frac{\delta(2\eta-1)(2\eta-2)}{2[\rho + \delta(1-\eta)]} < 0 \tag{21}$$

holds with $0 < \eta < 0.5$. The second term on the left is strictly convex approaching zero when $\eta \to 0.5$, and approaching $\delta/(\rho+\delta) > 0$ when $\eta \to 0$, so its supremum is $\delta/(\rho+\delta)$ over $\eta = 0$. The first term is strictly decreasing in η, and its supremum is $\ln(\rho+\delta)$ over $\eta = 0$. As each term has its supremum over $\eta = 0$, the supremum of the whole expression is here. Consequently, a sufficient condition for (21) to hold is the inequality

$$f(\delta, \rho) \equiv \ln[\rho + \delta] + \frac{\delta}{\rho + \delta} < 0 \tag{22}$$

in which $f(\delta, \rho)$ is strictly increasing in both variables and approaches $-\infty$ when $\delta \to 0$ and $\rho \to 0$. The next figure aids in understanding some further basic characteristics of the function in (22). In the right pane, the curve $f(\delta, \rho) = 0$ halves the (δ, ρ) space with positive values of $f(\delta, \rho)$ lying on the right.

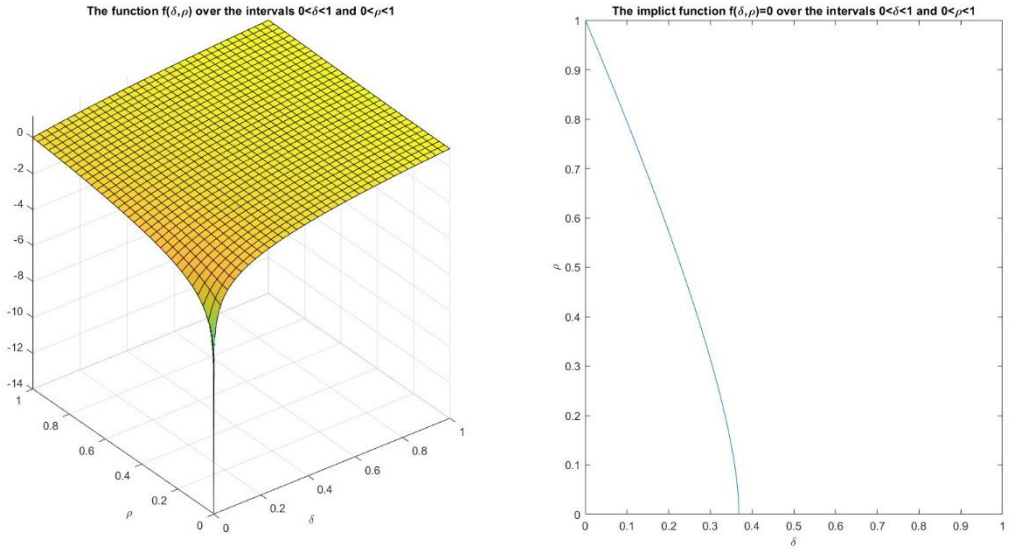


Fig. 5: The function $f(\delta, \rho)$ over the intervals $0 < \delta < 1$ and $0 < \rho < 1$ (left pane), and the implicit function $f(\delta, \rho) = 0$ over the same intervals (right pane).

As the figure suggests, the inequality in (22) holds for a wide range of normal values for δ at normally low levels of ρ. Due to sufficiency, (21) holds for a wider range of parameters than for which (22) holds. In the second term for θ^*, direct differentiation shows the exponent $0.5 < (2 - 2\eta)^{-1} < 1$ to be an increasing function of η, so the whole term is a negative function of η as long as $A^2 > \delta c$. However, for high values of c, $\theta_\eta^* > 0$ *might* occur since the monotonicity of θ_η^* depends on the monotonicity of the two terms taken together—even if the first term drops in response to an increase in η, the second term might offset it by a sharper increase.

By the same direct calculation of the differentials of h^* (Appendix A) for the $0 < \eta < 0.5$ case, we also have $h_c^* < 0$, so higher subjective costs prevent the representative agent from fulfilling her potential in learning, as we expected at the outset. No matter how quickly a student could learn, suffering from the daily chores of mugging at home discourages her from learning. The table below summarizes the first-order partial differentials of the steady-state values with respect to the parameters.

Table 1: Partial differentials of h^* and θ^* with respect to the parameters of the problem.

$h_\eta^* > 0$	$\theta_\eta^* = +/-$
$h_A^* > 0$	$\theta_A^* < 0$
$h_c^* < 0$	$\theta_c^* > 0$
$h_\delta^* < 0$	$\theta_\delta^* > 0$

Therefore, the equilibrium allocation to which the system converges as a sequence of temporary equilibria moves to the bottom right as $h_A^* > 0$ and $\theta_A^* < 0$, while the sign of θ_η^* is uncertain, so the effect on the endpoint of a rise in η due to the additive increase from γ in the social setting is not clear in cases where c is 'extremely' high. For 'normally' high values of c we can expect $\theta_\eta^* < 0$ to hold. The stability of the equilibrium dynamics directly follows from the stability of the dynamics of the isolated individual: At any point in time and at any exogenously given human capital of the classmate, the individual's capital stock converges toward a stable capital stock (Appendix H).

We can find the optimum allocation with the external effect by solving the Hamiltonian

$$\widetilde{H}(h_1, h_2, \theta_1, \theta_2, u_1, u_2) = \ln(h_1) + \ln(h_2) - \frac{c}{2}(u_1)^2 - \frac{c}{2}(u_2)^2 + \theta_1\left(Ah_2^{\frac{\gamma}{2}}h_1^{\eta_1+\frac{\gamma}{2}}u_1 - \delta h_1\right) + \theta_2\left(Ah_1^{\frac{\gamma}{2}}h_2^{\eta_2+\frac{\gamma}{2}}u_2 - \delta h_2\right). \tag{23}$$

Just like previously, when $\gamma = 0$, human capital accumulation is not affected by the knowledge of the classmate. In this case, the two problems are not related, and the social planner solves two separate problems in isolation, and we get back to the case of (a simple sum of) isolated problems. Following the standard procedure, we receive a six-equation system for the necessary conditions:

$$u_1 = \frac{A}{c}\theta_1 h_2^{\frac{\gamma}{2}}h_1^{\eta_1+\frac{\gamma}{2}} \tag{24a}$$

$$u_2 = \frac{A}{c}\theta_2 h_1^{\frac{\gamma}{2}}h_2^{\eta_2+\frac{\gamma}{2}} \tag{24b}$$

$$\dot{\theta}_1 = (\rho+\delta)\theta_1 - \frac{1}{h_1} - A\left(\eta_1 + \frac{\gamma}{2}\right)\theta_1 h_1^{\eta_1+\frac{\gamma}{2}-1}h_2^{\frac{\gamma}{2}}u_1 - A\frac{\gamma}{2}\theta_2 h_1^{\frac{\gamma}{2}-1}h_2^{\eta_2+\frac{\gamma}{2}}u_2 \tag{25a}$$

$$\dot{\theta}_2 = (\rho+\delta)\theta_2 - \frac{1}{h_2} - A\left(\eta_2 + \frac{\gamma}{2}\right)\theta_2 h_1^{\frac{\gamma}{2}}h_2^{\eta_2+\frac{\gamma}{2}-1}u_2 - A\frac{\gamma}{2}\theta_1 h_1^{\eta_1+\frac{\gamma}{2}}h_2^{\frac{\gamma}{2}-1}u_1 \tag{25b}$$

$$\dot{h}_1 = Ah_2^{\frac{\gamma}{2}}h_1^{\eta_1+\frac{\gamma}{2}}u_1 - \delta h_1 \tag{26a}$$

$$\dot{h}_2 = Ah_1^{\frac{\gamma}{2}}h_2^{\eta_2+\frac{\gamma}{2}}u_2 - \delta h_2 \tag{26b}$$

that can be used to find h_1^* and h_2^* as the optimum levels of knowledge for the two families, and θ_1^* and θ_2^* as the optimum values of the costate variables, all emerging where $\dot{h}_1 = \dot{h}_2 = \dot{\theta}_1 = \dot{\theta}_2 = 0$. The first-order conditions for $u_1(t)$ and $u_2(t)$ established in (24a) and (24b) can be substituted into the equations of the states and the costates that are all zero on the left in the steady state. By doing so, we have a system of four equations in four variables (h_1, h_2, θ_1 and θ_2). Then, having taken $\dot{h}_1 = \dot{h}_2 = 0$ in (26a) and (26b), we have two equations that can be solved for θ_1 and θ_2 as

$$\theta_1 = \frac{\delta c}{A^2} h_1^{1-2\eta_1-\gamma} h_2^{-\gamma} \tag{27a}$$

$$\theta_2 = \frac{\delta c}{A^2} h_1^{-\gamma} h_2^{1-2\eta_2-\gamma} \tag{27b}$$

that can be substituted into (25a) and (25b). With this step, we simplify the above system of six equations to a system of two equations in h_1 and h_2. By working simultaneously with (25a) and (25b) and using the prevalent symmetries in the system, we receive two separate equations for $2A^2 h_1^{\gamma} h_2^{\gamma}$ that, taken together, imply

$$h_2^* = \left[\frac{2\left[\delta\rho c + \delta^2 c\left(1-\eta_1-\frac{\gamma}{2}\right)\right] + \delta^2\gamma c}{2\left[\delta\rho c + \delta^2 c\left(1-\eta_2-\frac{\gamma}{2}\right)\right] + \delta^2\gamma c}\right]^{\frac{1}{2-2\eta_2}} (h_1^*)^{\frac{1-\eta_1}{1-\eta_2}} \tag{28}$$

that is an Euler equation clarifying the optimum ratio of the state variables. Note that the case $\eta_1 = \eta_2$ implies $h_1^* = h_2^*$ by (28). In the case when individuals are alike hence $\eta_1 = \eta_2 \equiv \eta$, we have $h_1(t) = h_2(t)$ for all t, and in the optimal allocation $0 < \eta + \gamma < 0.5$ must hold, under which $0 < \eta_j + \frac{\gamma}{2} < 0.5$ obviously holds. The previous assumption $\eta_1 > \eta_2$ means that in (28) we have

$$0 < \frac{2\left[\delta\rho c + \delta^2 c\left(1-\eta_1-\frac{\gamma}{2}\right)\right] + \delta^2\gamma c}{2\left[\delta\rho c + \delta^2 c\left(1-\eta_2-\frac{\gamma}{2}\right)\right] + \delta^2\gamma c} < 1,\ \frac{1}{2} < \frac{1}{2-2\eta_2} < 1, \text{ and } 0 < \frac{1-\eta_1}{1-\eta_2} < 1$$

with $h_1^* > 1$, so eventually $h_2^* < h_1^*$. The optimum level of human capital for the family with lower learning abilities falls short of the optimum knowledge level of the other family.

Going back to the previous step in which, taken $\dot{\theta}_1 = \dot{\theta}_2 = 0$, we have two equations in h_1 and h_2, we can use (28) to find that

$$\begin{aligned} 1 &= \left\{\frac{\delta\rho c + \delta^2 c\left(1-\eta_1-\frac{\gamma}{2}\right)}{A^2}\left[\frac{2\left[\delta\rho c + \delta^2 c\left(1-\eta_2-\frac{\gamma}{2}\right)\right] + \delta^2\gamma c}{2\left[\delta\rho c + \delta^2 c\left(1-\eta_1-\frac{\gamma}{2}\right)\right] + \delta^2\gamma c}\right]^{\frac{\gamma}{2-2\eta_2}}\right. \\ &\left. - \frac{\delta^2\gamma c}{2A^2}\left[\frac{2\left[\delta\rho c + \delta^2 c\left(1-\eta_1-\frac{\gamma}{2}\right)\right] + \delta^2\gamma c}{2\left[\delta\rho c + \delta^2 c\left(1-\eta_2-\frac{\gamma}{2}\right)\right] + \delta^2\gamma c}\right]^{\frac{2-2\eta_2-\gamma}{2-2\eta_2}}\right\} (h_1^*)^{\frac{(2-\gamma)(1-\eta_1-\eta_2)-\gamma+2\eta_1\eta_2}{1-\eta_2}} \\ &\equiv h(h_1^*) \end{aligned} \tag{29}$$

which uniquely determines h_1^* that positively depends on η_2 (Appendix I). Intuitively, when learning is a social activity with interaction of agents, the student with inferior learning abilities can hold her fellow classmate back in fulfilling her learning potential.

Most importantly, (28) and (29) describe h_1^* and h_2^* as functions of η_1, and η_2, just like in the equilibrium case. We expect the same result to hold in the optimum. It can be formally demonstrated that each optimum knowledge level is positively influenced by both η_1 and η_2. These results align with the intuition that the optimum capital stocks for both agents in the social setting in charge of the social planner are increasing functions of their learning abilities.

There is thus a unique h_1^* and by (28) a unique h_2^* to which there belong a pair of unique θ_1^* and θ_2^* by (27a) and (27b) as closed-form solutions to which the system should converge under the control by the social planner in the extension of the model of the isolated individual. As a final step at this stage of the analysis, we must examine if the system of two students exhibits saddle-path stability by checking if the optimum is stable in a sufficiently large neighborhood of the steady state. As the system now has two states and two costates, a qualitative phase diagram analysis is useless. The Jacobian of the system linearized around the steady state is

$$J = \begin{bmatrix} J_{11} & J_{12} & J_{13} & J_{14} \\ J_{21} & J_{22} & J_{23} & J_{24} \\ J_{31} & J_{32} & J_{33} & J_{34} \\ J_{41} & J_{42} & J_{43} & J_{44} \end{bmatrix}$$

where

$$J_{11} = \left(\dot{h}_1\right)_{h_1} = \frac{A^2}{c}(2\eta_1 + \gamma)\theta_1 h_1^{2\eta_1+\gamma-1} h_2^{\gamma} - \delta$$

$$J_{12} = \left(\dot{h}_1\right)_{h_2} = \frac{\gamma A^2}{c}\theta_1 h_1^{2\eta_1+\gamma} h_2^{\gamma-1}$$

$$J_{13} = \left(\dot{h}_1\right)_{\theta_1} = \frac{A^2}{c} h_1^{2\eta_1+\gamma} h_2^{\gamma}$$

$$J_{14} = \left(\dot{h}_1\right)_{\theta_2} = 0$$

$$J_{21} = \left(\dot{h}_2\right)_{h_1} = \frac{\gamma A^2}{c}\theta_2 h_1^{\gamma-1} h_2^{2\eta_2+\gamma}$$

$$J_{22} = \left(\dot{h}_2\right)_{h_2} = \frac{A^2}{c}(2\eta_2 + \gamma)\theta_2 h_1^{\gamma} h_2^{2\eta_2+\gamma-1} - \delta$$

$$J_{23} = \left(\dot{h}_2\right)_{\theta_1} = 0$$

$$J_{24} = \left(\dot{h}_2\right)_{\theta_2} = \frac{A^2}{c} h_1^{\gamma} h_2^{2\eta_2+\gamma}$$

$$J_{31} = \left(\dot{\theta}_1\right)_{h_1} = h_1^{-2} - \frac{A^2}{c}\left(\eta_1 + \frac{\gamma}{2}\right)(2\eta_1 + \gamma - 1)\theta_1^2 h_1^{2\eta_1+\gamma-2} h_2^{\gamma} - \frac{\gamma A^2}{2c}(\gamma - 1)\theta_2^2 h_1^{\gamma-2} h_2^{2\eta_2+\gamma}$$

$$J_{32} = \left(\dot{\theta}_1\right)_{h_2} = -\frac{\gamma A^2}{c}\left(\eta_1 + \frac{\gamma}{2}\right)\theta_1^2 h_1^{2\eta_1+\gamma-1} h_2^{\gamma-1} - \frac{\gamma A^2}{2c}(2\eta_2 + \gamma)\theta_2^2 h_1^{\gamma-1} h_2^{2\eta_2+\gamma-1}$$

$$J_{33} = \left(\dot{\theta}_1\right)_{\theta_1} = (\rho + \delta) - \frac{2A^2}{c}\left(\eta_1 + \frac{\gamma}{2}\right)\theta_1 h_1^{2\eta_1+\gamma-1} h_2^{\gamma}$$

$$J_{34} = \left(\dot{\theta}_1\right)_{\theta_2} = -\frac{\gamma A^2}{c}\theta_2 h_1^{\gamma-1} h_2^{2\eta_2+\gamma}$$

$$J_{41} = \left(\dot{\theta}_2\right)_{h_1} = -\frac{\gamma A^2}{c}\left(\eta_2 + \frac{\gamma}{2}\right)\theta_2^2 h_1^{\gamma-1} h_2^{2\eta_2+\gamma-1} - \frac{\gamma A^2}{2c}(2\eta_1+\gamma)\theta_1^2 h_1^{2\eta_1+\gamma-1} h_2^{\gamma-1}$$

$$J_{42} = \left(\dot{\theta}_2\right)_{h_2} = h_2^{-2} - \frac{A^2}{c}\left(\eta_2 + \frac{\gamma}{2}\right)(2\eta_2+\gamma-1)\theta_2^2 h_1^{\gamma} h_2^{2\eta_2+\gamma-2} - \frac{\gamma A^2}{2c}(\gamma-1)\theta_1^2 h_1^{2\eta_1+\gamma} h_2^{\gamma-2}$$

$$J_{43} = \left(\dot{\theta}_2\right)_{\theta_1} = -\frac{\gamma A^2}{c}\theta_1 h_1^{2\eta_1+\gamma} h_2^{\gamma-1}$$

$$J_{44} = \left(\dot{\theta}_2\right)_{\theta_2} = (\rho+\delta) - \frac{2A^2}{c}\left(\eta_2 + \frac{\gamma}{2}\right)\theta_2 h_1^{\gamma} h_2^{2\eta_2+\gamma-1}$$

to be evaluated in the steady state with h_1^*, h_2^*, θ_1^* and θ_2^* given by (29), (28), (27a) and (27b). The eigenvalues of the Jacobian are not particularly informative in their closed forms (even if they do exist), so we must make do with the claim that under a wide array of normal parameter configurations the linearized system proves to have two positive and two negative eigenvalues. From this, in each case, a slope matrix with all negative entries follow that implies that when the states are under their steady-state values, the costates are above their corresponding optimum values, so, just like in the isolated case, we can expect $\dot{h}_i(t) > 0$ and $\dot{\theta}_i(t) < 0$, $i = 1,2$. Even though the introduction of the external effect modifies the equilibrium, it neither changes the uniqueness of the optimum nor destabilize dynamics of the system that still proves to show saddle-path stability.

7. Comparative dynamics

As we have seen above, learning in a social setting affects the learning parameters A and η of the isolated case. To trace back to the isolated case how the individuals behave in a social environment, we need to understand how changes in A (taken as $Ah_i^{\frac{\gamma}{2}}$ in the learning function for each agent) and η changes the behavior of the isolated system. The former change comes from the increase of the capital stock of the fellow classmate, while the latter stems from how the external effect elevates the learning parameters η_i, $i = 1,2$. The endpoints of the convergence in the isolated case are pinned down by h^* and θ^* the behavior of which is revealed by the partial differentials of the optimum solutions of the isolated case. However, these results are informative about the endpoints of convergence only. To see how the optimal trajectories change in the equilibrium and the optimum of the social case, we need to examine the comparative dynamics of the isolated case. Therefore, we still reduce the dynamics of the social setting to the case of the isolated individual as much as possible. To do so, we combine Oniki's (1973) and Caputo's (1990a) phase diagram approach with Caputo's (1990b; 2005, pp. 287-311) dynamic envelope theorem.

The qualitative, phase diagram approach is built on the idea that a change in a parameter modifies $\dot{h}$ and $\dot{\theta}$ through changing the paths of h and θ as well as, with given h_0, the starting value $\theta(0)$

along the saddle path. Most of the work is to be done in order to find the starting points and endpoints of the perturbed system. We can simply skip these steps as, with given h_0, we know from the isoquant analysis that $\theta(0)_\eta > 0$ and $\theta(0)_A < 0$, and that how changes in η and A affect the endpoints h^* and θ^* (note that the sign of θ^*_η is uncertain). So, the only step to take in order to draw the phase diagrams is to differentiate the system of law of motions for h and θ in (10) and (11) with respect to η and A (Appendix A). The differentiation with respect to η yields the equations

$$\dot{h}_\eta = \frac{A^2}{c} h^{2\eta} \theta_\eta + \left(\frac{2\eta A^2}{c} \theta h^{2\eta-1} - \delta \right) h_\eta + \frac{2A^2}{c} \theta h^{2\eta} \ln(h)$$

$$\dot{\theta}_\eta = \left[(\rho + \delta) - \frac{2\eta A^2}{c} \theta h^{2\eta-1} \right] \theta_\eta + \left[h^{-2} + \frac{\eta A^2}{c} (1 - 2\eta) \theta^2 h^{2\eta-2} \right] h_\eta - \frac{A^2}{c} \theta^2 h^{2\eta-1} - \frac{\eta A^2}{c} \theta^2 [2h^{2\eta-1} \ln(h)]$$

that takes the general form

$$\begin{bmatrix} \dot{\theta}_\eta \\ \dot{h}_\eta \end{bmatrix} = \begin{bmatrix} +/- & + \\ + & +/- \end{bmatrix} \begin{bmatrix} \theta_\eta \\ h_\eta \end{bmatrix} + \begin{bmatrix} - \\ + \end{bmatrix}$$

from which the phase diagram

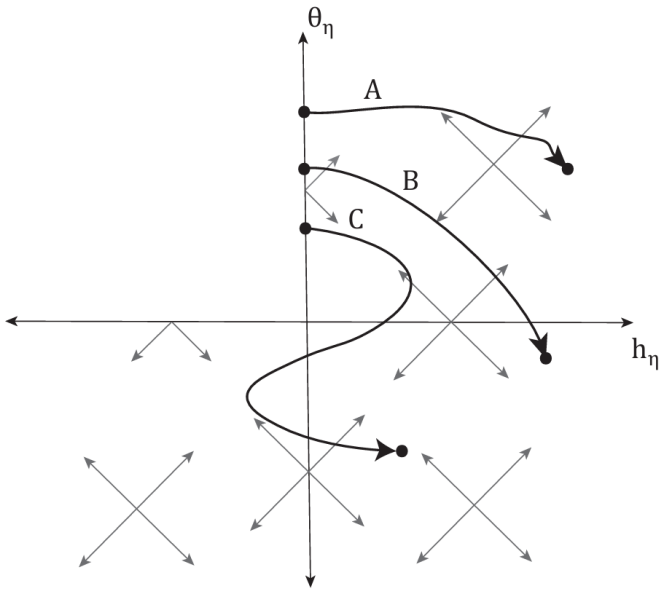


Fig. 6: Phase diagram of the convergence perturbed by an increase in η.

follows. As $\theta(0)_\eta > 0$ and with a given h_0, after an increase in η the perturbed process starts somewhere on the positive segment of the vertical axis from which no route leads to the left. As we expect $h^*_\eta > 0$ while θ^*_η is uncertain, the endpoints of the perturbed convergences must lie in the positive half of the plane lying to the right from the vertical axis. Paths A and B are such

typical cases. From the starting points, we move to the right that means that an increase in η elevates the whole capital profile, so the change leads to higher capital stocks at any t. However, based on the directions of flows, we cannot exclude a case such as path C a part of which lies to the left of the vertical axis meaning that while $\dot{h} > 0$ holds all along, an increase in η *lowers* the capital profile and slows the speed of the accumulation of the capital stock for a while. In such a case, the continuously increasing capital profile with a higher η is temporarily *under* the capital profile implied by a lower η. We shall return to this case below.

The differentiation with respect to A yields the equations

$$\dot{h}_A = \frac{A^2}{c} h^{2\eta} \theta_A + \left(\frac{2\eta A^2}{c} \theta h^{2\eta-1} - \delta \right) h_A + \frac{2A}{c} \theta h^{2\eta}$$

$$\dot{\theta}_\eta = \left[(\rho + \delta) - \frac{2\eta A^2}{c} \theta h^{2\eta-1} \right] \theta_A + \left[h^{-2} + \frac{\eta A^2}{c} (1 - 2\eta) \theta^2 h^{2\eta-2} \right] h_A - \frac{2\eta A}{c} \theta^2 h^{2\eta-1}$$

that takes the general form

$$\begin{bmatrix} \dot{\theta}_A \\ \dot{h}_A \end{bmatrix} = \begin{bmatrix} +/- & + \\ + & +/- \end{bmatrix} \begin{bmatrix} \theta_A \\ h_A \end{bmatrix} + \begin{bmatrix} - \\ + \end{bmatrix}$$

from which the phase diagram follows

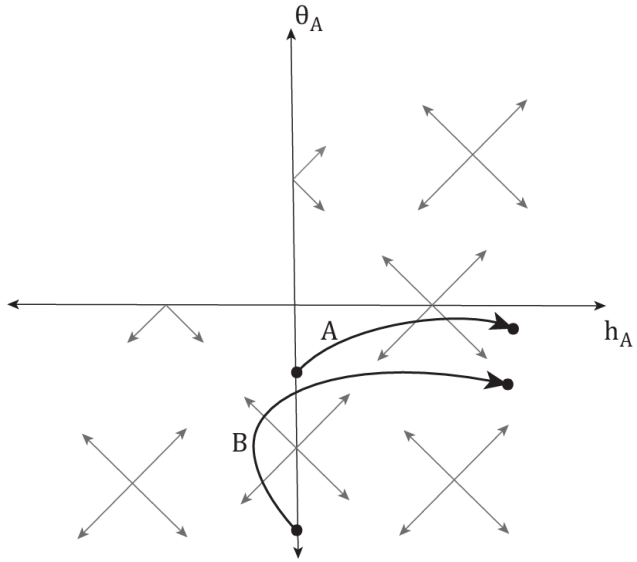


Fig. 7: Phase diagram of the convergence perturbed by an increase in *A*.

As $\theta(0)_A < 0$ and with a given h_0, after an increase in A, the perturbed process starts somewhere on the negative segment of the vertical axis from which routes do lead to the left as well. It is also known that $h_A^* > 0$ and $\theta_A^* < 0$, so the endpoints certainly lie in the bottom right quadrant. Path A is a typical case where the whole convergence process runs on the right of the vertical axis

meaning that an increase in A elevates the whole capital profile. In this case, a higher A, other things equal, implies higher capital stock at any $t > 0$ (note that h_0 is fixed). However, just like in the perturbation analysis of η, a case like path B cannot be ruled out, so it might happen that an increase in A temporarily slows the speed of capital accumulation. In such a case, while $\dot{h} > 0$ holds all along the accumulation process, an initial segment of the capital profile implied by a higher A is *under* the capital path implied by a lower A.

These results can be revisited and sharpened by Caputo's envelope theorem framework. As there is a constraint on the control given as $0 \leq u(t) \leq 1$, the Lagrangian of the isolated case is simply

$$L = \ln(h) - \frac{c}{2}u^2 + \theta(Ah^{\eta}u - \delta h) + \mu(u - 1)$$

in which μ is the costate of the Kuhn–Tucker expression corresponding to the control variable. The dynamic envelope theorem results come from the differentiation of L with respect to the relevant parameters, now η and A, then evaluate and integrate the expression along the optimal paths $h^*(t,\beta)$, $\theta^*(t,\beta)$ and $u^*(t,\beta)$. As the Kuhn–Tucker part does not depend on either η or A, we can simply differentiate the Hamiltonian. The results

$$\partial L/\partial A = \theta h^{\eta}u$$

$$\partial L/\partial A \;|_{\text{optimal path}} = \theta^*(t,\beta)\,[h^*(t,\beta)]^{\eta}\,u^*(t,\beta)$$

$$J_A^*(\beta) = \int_0^{\infty} \theta^*(t,\beta)\,[h^*(t,\beta)]^{\eta}\,u^*(t,\beta)\,dt > 0 \tag{30}$$

follow, in which the parameter β of givens contains not only A, c, η, and δ, but also h_0. Given that $1 < h_0 < h^*$ applies by assumption, the analogue result with respect to η is

$$J_{\eta}^*(\beta) = A\int_0^{\infty} \theta^*(t,\beta)\,[h^*(t,\beta)]^{\eta}\,\ln[h^*(t,\beta)]\,u^*(t,\beta)\,dt > 0 \tag{31}$$

according to which a rise in either η or A increases the value to be obtained with the optimal behavior. In the next step, (30) and (31) are to be differentiated with respect to η and A that yield

$$H_{11} \equiv J_{AA}^*(\beta) = \int_0^{\infty} \frac{\partial}{\partial A}\theta^*(t,\beta)\,[h^*(t,\beta)]^{\eta}\,u^*(t,\beta)\,dt = 0$$

$$H_{22} \equiv J_{\eta\eta}^*(\beta) = A\int_0^{\infty} \frac{\partial}{\partial \eta}\theta^*(t,\beta)\,[h^*(t,\beta)]^{\eta}\,\ln[h^*(t,\beta)]\,u^*(t,\beta)\,dt$$

$$= A\int_0^{\infty} \theta^*(t,\beta)\,[h^*(t,\beta)]^{\eta}\,\ln[h^*(t,\beta)]\,\ln[h^*(t,\beta)]\,u^*(t,\beta)\,dt > 0$$

$$H_{12} = H_{21} \equiv J^*_{A\eta}(\beta) = \int_0^\infty \frac{\partial}{\partial \eta}\theta^*(t,\beta)\,[h^*(t,\beta)]^\eta\, u^*(t,\beta)\,dt$$
$$= \int_0^\infty \theta^*(t,\beta)\,[h^*(t,\beta)]^\eta\, \ln[h^*(t,\beta)]\, u^*(t,\beta)\,dt = J^*_{\eta A}(\beta) > 0$$

that can be summarized in the Hessian

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} = \begin{bmatrix} 0 & + \\ + & + \end{bmatrix}.$$

Given the general structure of H, its determinant $|H| < 0$. Given that the determinant is the product of the eigenvalues, the eigenvalues have opposite signs as both eigenvalues are real (Horn & Johnson, 2013, pp. 50-51). Using the fact that the eigenvalues of a positive semidefinite matrix are nonnegative, and the eigenvalues of a negative semidefinite matrix are nonpositive with strict inequalities for definite matrices (Horn & Johnson, 2013, p. 430), the mixed signs imply that H is indefinite.[3]

Note, however, that when η is treated on its own, judged by the positive sign in the $(2,2)$ position of H, we have a one-element Hessian $H_{22} \in \mathbb{R}^{1\times 1}$ which is positive definite. Therefore, as the integral in (31) is strictly positive, a marginal rise in η increases the value of the optimal behavior, and through the optimality conditions results in a higher optimal capital profile $h^*(t,\beta)$ starting from any $1 < h_0 < h^*$ under any appropriate parameter configuration, for all $t > 0$ and with no change in other parameters. Consequently, $J^*(\beta)$ is a strictly increasing and strictly (locally) convex function of η, so the effect of η on the state profile might even be accelerated by further marginal increases, at least these effects are not to be expected to fade away. This result rules out path C in Fig. 6.

We have a similar result for A. As the integral in (30) is strictly positive, a marginal rise in A increases the value of the optimal behavior, and through the optimality conditions results in a higher optimal capital profile $h^*(t,\beta)$ starting from any $1 < h_0 < h^*$ under any appropriate parameter configuration, for all $t > 0$ and all other parameters kept fixed, as the Hessian $H_{11} \in \mathbb{R}^{1\times 1}$ of A is zero hence semidefinite. Consequently, $J^*(\beta)$ is a strictly increasing and linear function of A. This result excludes path B in Fig. 7.

---

[3] If we treat all the four parameters A, η, c, and δ, the Hessian

$$\widehat{H} = \begin{bmatrix} 0 & + & 0 & 0 \\ + & + & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

is still indefinite. To see this, we just need to use the fact that $\widehat{H}$ is a block diagonal matrix, and the eigenvalues of $\widehat{H}$ are the eigenvalues of the top left and bottom right blocks, the bottom left and top right blocks being zero. While the bottom right block has two zero eigenvalues, the top left block has just proven to be indefinite. In Caputo's framework, the direction of the effect of a parameter change is determined by the integrated differential of the Lagrangian. To have clear results, the Hessian must be either positive or negative definite or semi-definite.

The fact that H is indefinite implies that $J^*(\beta)$ is neither a strictly convex nor a strictly concave function of the parameters η and A. Therefore, there might be some direction in the two-dimension parameter space $\eta \times A \subset \mathbb{R}^2$ along which increasing η and A simultaneously might lead to ambiguous or even counteracting effects on the capital profile of the isolated individual. It means that when increasing η and A simultaneously under a parameter configuration, the atypical behavior of a temporary capital profile drop *might* occur. In this case, given any $h(t)$, keeping $\dot{h}$ above zero is possible with lower $u(t)$, so the atypical case is straightforward to interpret as saving learning time due to higher productivity in learning.

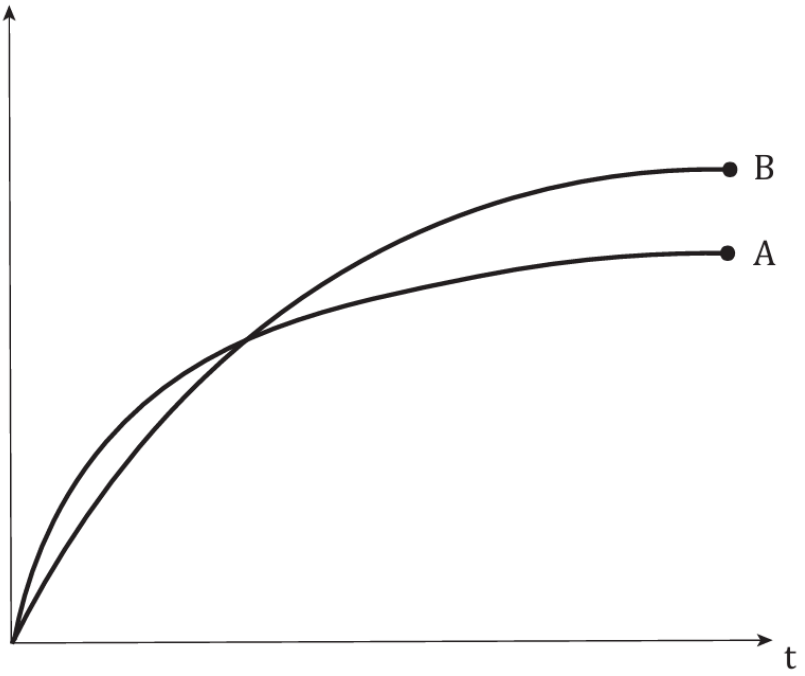


Fig. 8: A schematic representation of the atypical behavior of saving learning time in the isolated case. The baseline case is *A*, while the path running toward *B* takes place after a simultaneous increase in η and *A*.

When applying these results to the equilibrium allocation and paths, however, we must consider the fact that the parameter $\eta_i + \gamma/2$ is fixed for both agents $(i = 1,2)$ all over the capital accumulation process. Therefore, simultaneous changes in $\eta_i + \gamma/2$ and A taken as changes in $Ah_j^{\frac{\gamma}{2}}$ are excluded, only $Ah_j^{\frac{\gamma}{2}}$ is increasing in time as $h_j(t)$ is increasing in time $(j = 1,2)$. This change, however, can easily traced back to changes in A of the isolated case above.

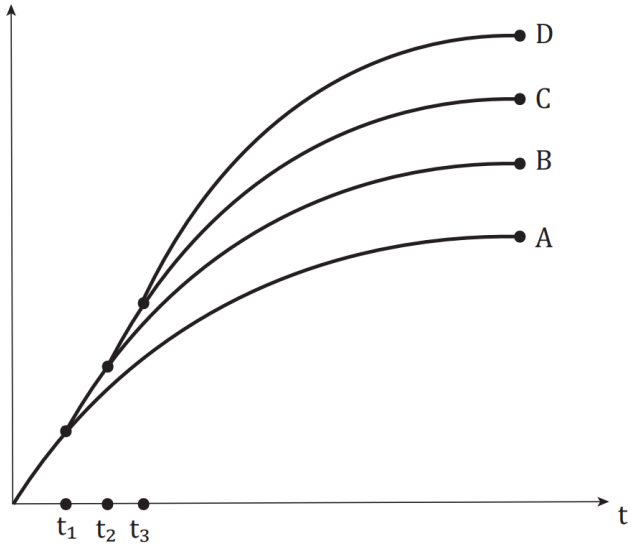


Fig. 9: A schematic representation of the capital profile of the individual in the equilibrium setting. The actual path is the upper envelope curve of the series of paths. The temporary endpoints *A*, *B*, *C* and *D* are continuously updated and increasing as the capital stock of the fellow classmate is increasing. This continuous update is approximated with discrete-time updates taking place in t_1, t_2 and t_3.

In every moment, the increase in the capital stock of the fellow classmate opens a new capital path for the individual in the equilibrium setting that she experiences as an exogenous increase in parameter A in the learning function. Under any appropriate parameter configuration a part of which is the capital stock of the fellow classmate, there is a stable manifold along which $\dot{h}_i > 0$ holds. The saddle-path property seen above guarantees local uniqueness and convergence at every moment and at any capital distribution in the setting. The increase in the capital stock of the fellow classmate at any moment leads to a higher capital path in line with how the endpoints of capital convergence are increasing. Given the results from comparative dynamics above, to our previous insights we can add the additional finding that to any temporary endpoint of capital convergence (points A, B, C and D in Fig. 9) there belongs a unique capital path for each agent that is *above* all the capital paths belonging to any previous temporary endpoints.

These results are easy to extend to the optimum of the social setting defined in (16) and (23). In this two-state and two-costate case we cannot use the phase diagram approach, so we must start out with the application of the dynamic envelope theorem. The Lagrangian of the problem is

$$\tilde{L} = \tilde{H}(h_1, h_2, \theta_1, \theta_2, u_1, u_2) = \ln(h_1) + \ln(h_2) - \frac{c}{2}(u_1)^2 - \frac{c}{2}(u_2)^2 + \theta_1\left(Ah_2^{\frac{\gamma}{2}}h_1^{\eta_1+\frac{\gamma}{2}}u_1 - \delta h_1\right) +$$
$$\theta_2\left(Ah_1^{\frac{\gamma}{2}}h_2^{\eta_2+\frac{\gamma}{2}}u_2 - \delta h_2\right) + \mu_1(u_1 - 1) + \mu_2(u_2 - 1).$$

the direct differentiation of which yields

$$\partial\tilde{L}/\partial A = \theta_1 h_2^{\frac{\gamma}{2}}h_1^{\eta_1+\frac{\gamma}{2}}u_1 + \theta_2 h_1^{\frac{\gamma}{2}}h_2^{\eta_2+\frac{\gamma}{2}}u_2$$

from which the integration along the optimum paths leads to

$$\tilde{J}_A^*(\beta) = \int_0^\infty \theta_1^*(t,\beta)\,[h_2^*(t,\beta)]^{\frac{\gamma}{2}}\,[h_1^*(t,\beta)]^{\eta_1+\frac{\gamma}{2}}\,u_1^*(t,\beta) + \theta_2^*(t,\beta)\,[h_1^*(t,\beta)]^{\frac{\gamma}{2}}\,[h_2^*(t,\beta)]^{\eta_2+\frac{\gamma}{2}}\,u_2^*(t,\beta)\,dt$$
$$> 0$$

that is certainly positive. Of the other two partial differentials, the first one is

$$\partial\tilde{L}/\partial\eta_1 = \theta_1 A h_2^{\frac{\gamma}{2}} h_1^{\eta_1+\frac{\gamma}{2}} \ln(h_1)\,u_1$$

from which the integral along the optimum paths

$$\tilde{J}_{\eta_1}^*(\beta) > 0$$

is certainly positive as long as the starting value $h_1(0) > 1$ holds by assumption, as an analogue of the starting value in the isolated case. By the basic symmetry of the problem

$$\tilde{J}_{\eta_2}^*(\beta) > 0$$

also follows. Using the linearity of $\tilde{J}^*(\beta)$ in parameter A, we have the upper left element in the Hessian $\widetilde{H}$ of the problem as $\widetilde{H}_{11} \equiv \tilde{J}_{AA}^*(\beta) = 0$. Most importantly,

$$\widetilde{H}_{22} \equiv \tilde{J}_{\eta_1\eta_1}^*(\beta) = A\int_0^\infty \frac{\partial}{\partial\eta_1}\,\theta_1^*(t,\beta)\,[h_2^*(t,\beta)]^{\frac{\gamma}{2}}\,[h_1^*(t,\beta)]^{\eta_1+\frac{\gamma}{2}} \ln[h_1^*(t,\beta)]\,u_1^*(t,\beta)\;dt$$
$$= A\int_0^\infty \theta_1^*(t,\beta)\,[h_2^*(t,\beta)]^{\frac{\gamma}{2}}\,[h_1^*(t,\beta)]^{\eta_1+\frac{\gamma}{2}}\,\ln[h_1^*(t,\beta)]\ln[h_1^*(t,\beta)]\;u_1^*(t,\beta)\;\;dt > 0$$

under the previous assumption regarding the starting value for the process $h_1(t)$. By the previous symmetry,

$$\widetilde{H}_{33} \equiv \tilde{J}_{\eta_2\eta_2}^*(\beta) > 0$$

also holds. Regarding the off-diagonal elements of the Hessian $\widetilde{H}$,

$$\widetilde{H}_{12} = \widetilde{H}_{21} \equiv \tilde{J}_{A\eta_1}^*(\beta) = \int_0^\infty \theta_1^*(t,\beta)\,[h_2^*(t,\beta)]^{\frac{\gamma}{2}}\,[h_1^*(t,\beta)]^{\eta_1+\frac{\gamma}{2}}\,\ln[h_1^*(t,\beta)]\;u_1^*(t,\beta) > 0$$

is certainly positive, and the symmetry of the problem also implies $\widetilde{H}_{13} = \widetilde{H}_{31} \equiv \tilde{J}_{A\eta_2}^*(\beta) > 0$. From $\partial\tilde{L}/\partial\eta_1$ and $\partial\tilde{L}/\partial\eta_2$ it follows that $\widetilde{H}_{23} = \widetilde{H}_{32} \equiv \tilde{J}_{\eta_1\eta_2}^*(\beta) = 0$. From this analysis, a Hessian

$$\widetilde{H} = \begin{bmatrix} 0 & + & + \\ + & + & 0 \\ + & 0 & + \end{bmatrix}$$

emerges in its general form. Again, $\widetilde{H}$ is indefinite (Appendix J). Accordingly, just like in the equilibrium setting, there is a direction in the parameter space $(\eta_1, \eta_2, A) \subset \mathbb{R}^3$ along which increasing η_1, η_2 and A simultaneously might lead to ambiguous or even counteracting effects on the capital profile of the isolated individual. However, it still holds that the parameter $\eta_i + \gamma/2$ is fixed for each individual ($i = 1,2$) all over the capital accumulation process. Therefore, only $Ah_i^{\frac{\gamma}{2}}$ is

increasing in time as $h_i(t)$ is increasing in time ($i = 1,2$), we thus receive the optimized dynamics of the paths depicted in Fig. 9 above.

8. Discussion: Rational segregation in a model of school learning

The dynamic envelope theorem results in (30) and (31) are easy to interpret as a drive for Agent 1 to select as high a parameter value for η_2 as possible. By itself, (31) shows that the way the agent's utility depends on η makes her interested in learning in a classroom environment instead of learning in isolation. The external effect of classroom learning adds to η. Beyond this interest, however, the actual benefit a student can realize from learning as a social activity crucially depends on parameter η of her fellow classmate. In the equilibrium setting, therefore, selecting a classroom environment boils down to selecting a parameter η_i of a possible classmate. Assuming $\eta_1 > \eta_2$, Agent 1 realizes that higher values for η_2 imply higher capital profiles for both agents. Given the starting values of the capital stocks, both agents in a classroom setting start converging toward their temporary endpoints being continuously increasing as their capital stocks are continuously increasing. Agent 1 as a utility-maximizer knows, however, that a higher η_2 leads to higher $h_2(t)$ at any $t > 0$ that also boosts her own capital accumulation process through (17). As overall discounted utility is an increasing function of A in the equilibrium setting by (30), a higher capital profile of Agent 2 makes Agent 1 happier and more satisfied, so Agent 1 is interested in finding a classmate with η_2 as close to η_1 possible and, equivalently, in avoiding agents with low η_2 values. The rational individual having high learning ability thus tends to segregate. The case is the exact opposite for Agent 2. The way her utility depends on η_1 of a possible classmate having higher learning abilities makes her interested in finding a classmate with a learning parameter as high as possible, which makes her interested in avoiding segregation. So, based on our previous results, we can claim the following theorems. As we can use all the previous results, we do not need a formal demonstration at this point.

Theorem 1: The asymmetric drive of the individual for segregation. *Assuming $\eta_1 > \eta_2$, both Student 1 and Student 2 are interested in finding a classmate with the highest possible learning parameter η_2 and η_1, respectively.*

Proof: Following from $(h_1^e)_{\eta_1} > 0$ and $(h_1^e)_{\eta_2} > 0$ we know how h_1^e and h_2^e depend on η_1 and η_2, yielding the result that each agent's utility is an increasing function of the classmate's learning parameter η_i (Proposition 6 and 7, Appendix G). So, we only need to show that:

(1) A higher learning parameter η_i of the classmate elevates the entire capital profile of the agent, and
(2) The rational utility maximizing agent prefers a higher capital profile, therefore, prefers a classmate having a higher learning parameter η_i to a classmate with a lower parameter value.

As for (1), due to the external effect, the classmate's continuously increasing capital stock is continuously increasing the effect of A for the agent, as shown in Fig. 9, placing the agent on a higher capital profile every moment. This part of the issue has already been resolved by the perturbation analysis in Section 7. However, a missing point is how η_i of the classmate impacts this process. The same perturbation analysis showed that a higher η_i elevates the entire own capital profile. From this, in the equilibrium setting it follows that at any t (with $t > t_0$) the classmate, thanks to her higher η_i, proceeds on a higher capital profile, as compared to any case with a lower η_i. This higher profile affects the agent's productivity in t through increasing the effect of her A in the learning function that leads to a higher capital profile as showed in Section 7. We also have to consider the fact that, thanks to her higher η_i, the classmate was on a higher capital profile in t as compared to the one that she would have been on with a lower learning parameter. Therefore, the classmate's higher η_i elevates her entire capital profile, and increases the effect of A in the learning function of the agent along her whole accumulation process. As a higher A elevates the entire capital profile for the agent, a higher η_i elevates the capital profiles for both agents in the equilibrium setting.

As for (2), given that the individual perceives the capital profile of the classmate through the external effect affecting A, the question regards if the utility maximizing individual is interested in a learning function with an increased A. Section 7 shows that a higher A increases the value of the optimal behavior, therefore, the individual is interested in having a higher A and, equivalently, in learning with a classmate who, thanks to her higher η_i, facilitates a higher A. As it holds for both agents in the equilibrium setting with $\eta_1 > \eta_2$, Agent 1 has a tendency to segregate and avoid any classmate with low η_2. By contrast, Agent 2 can maximize her utility with learning with a classmate having a η_1 as high as possible, so Agent 2 is against segregation.

□

We have a similar result in the optimum setting where time investments and capital paths are allocated by the social planner. To say something intelligent about the social planner's drive to segregate, however, we need to specify a mechanism for student selection. It makes a difference if the social planner maximizes total utility of the class with a given η_1 or η_2.

Theorem 2: The asymmetric drive of the social planner for segregation. *Assuming* $\eta_1 > \eta_2$*, given* η_i*, the social planner is interested in finding a classmate with the highest possible learning parameter* η_j*,* $i \neq j$.

Proof: As we have seen above, $\tilde{L}$ is an increasing function of η_1 and η_2. Still assuming $\eta_1 > \eta_2$ and with given η_1, the Hessian $\tilde{H}$ implies that the social planner can increase the overall utility of the class by selecting a classmate with parameter η_2 as close to η_1 as possible. In the opposite case, i.e. with given η_2, to maximize total utility the social planner picks a classmate with the highest possible η_1, which is a drive for setting up a heterogenous class.

This choice, at the end of the day, is the maximization of the maximum utility by an optimum choice of parameters. In both the equilibrium and the optimum settings, segregation is motivated by the idea that by selecting an appropriate classmate both the individual and the social planner can influence the parameters of human capital accumulation. In Theorem 2, we assumed that student selection is a process in which the social planner treats η_1 or η_2 as given, and his task is to find an appropriate classmate—this is what was regarded above as a mechanism of student selection.

□

The model as it stands now aligns with a number of empirical facts. At the most basic level, the model suggests that, other things equal, children from deprived families hence having low η_2 due to the lack of adequate home facilities and support naturally perform worse. As we have seen, a student with inferior learning abilities can benefit from a heterogenous or unsegregated class environment the most when she is mixed with one having very high η_1. Accordingly, the *higher* the difference between η_1 and η_2 while $\eta_1 > \eta_2$ holds, the more benefit is realized by Agent 2 in the form of an increase in the equilibrium or optimum capital stock, and in the elevation of the capital path. Similarly, top students benefit the most from the presence of similar students as the *smaller* the difference between η_1 and η_2 while $\eta_1 > \eta_2$ holds, the more benefit is realized by Agent 1 in the form of an increase in the equilibrium or optimum capital stock, and in the elevation of the capital path. The external effect in the model is asymmetric—the same heterogenous setting is harmful to the agent with high η_1 and beneficial to the agent with low η_2. The model thus predicts diverse responses to desegregation. Given that the model, other things equal, associates higher stocks of human capital to higher values of parameter η, it is straightforward to translate the implications of differences in η_i into differences in the levels of human capital. On this showing, the implications are consistent with the results by, for instance, Hoxby and Weingarth (2005) who found that students with low scores benefit the most from the addition of students from the top 15% of the score distribution. Likewise, top students benefit the most from the presence of students from their own league. Imberman, Kugler, and Sacerdote (2009) also found that high achieving students' school performance benefits the most from the presence or appearance of other high achieving students. Lavy, Silva, and Weinhardt (2012) report that all capability cohorts of students are harmed by the presence of the lowest achieving classmates.

The model also offers a theoretical explanation to why desegregation might fail. Desegregation efforts *without* strategies to improve the underlying factors such as home environments, facilities for early-childhood learning and family income hence to increase η_2 might not work. In such a case, desegregation means that priorly established peer networks are destroyed, so in a new, integrated environment a family with superior learning abilities is exposed to negative external effects delivered by families in low socio-economic status. This case aligns with the model's predictions according to which such a desegregation strategy is harmful to the family with high η_1

that ultimately leads to efforts to resegregate (Orfield & Lee, 2007) as low school performance in the model is not simply a consequence of segregation but can actively drive it. In this negative case, as the model predicts, students from disadvantageous backgrounds might find themselves on a downward spiral—low socioeconomic status implies poor home learning facilities that, in turn, hinder adequate progress in school that, in turn, intensifies segregation and resegregation. The benefits a family can realize with an increased η_2 are higher, and the policy measures increasing the parameter also dampen the efforts to resegregate.

Appendix A

Proposition 1: *The problem defined by (5) and (6) has a unique optimum steady-state solution* h^* *to which there belongs an optimal value* θ^* *for the costate variable. These solutions are finite at any configuration of appropriate and finite parameters.*

Proof: To find the steady state values, we should use the fact that the point $(h^*; \theta^*)$ in the space $S = (h, \theta)$ is pinned down by the intersection of the nullclines $\dot{h} = 0$ and $\dot{\theta} = 0$. By substituting $u = \frac{A}{c}\theta h^{\eta}$ from (4) into (11), the latter becomes

$$\dot{h} = \frac{A^2}{c}\theta h^{2\eta} - \delta h \tag{A.1}$$

that when $\dot{h} = 0$ leads to

$$\theta = \frac{\delta c}{A^2} h^{1-2\eta} \tag{A.2}$$

as the nullcline of h. Note that the shape of (A.2) depends on what value the parameter η takes. For $0 < \eta < 0.5$ it is strictly increasing, while for $0.5 < \eta < 1$ it is strictly decreasing. Finding the $\dot{\theta} = 0$ nullcline starts with the same substitution from (4) into (5) that becomes

$$\dot{\theta} = \theta\left(\rho + \delta - \frac{A^2\eta}{c}\theta h^{2\eta-1}\right) - \frac{1}{h} \tag{A.3}$$

that when $\dot{\theta} = 0$ gives the equation

$$(\rho + \delta) = \frac{1}{\theta h} + \frac{A^2\eta}{c}\theta h^{2\eta-1} \tag{A.4}$$

as the nullcline of θ. Equations (A.2) and (A.4) together imply that

$$\theta^* = \frac{\delta c}{A^2}(h^*)^{1-2\eta} \tag{A.5}$$

$$(\rho + \delta) = \frac{1}{\theta^* h^*} + \frac{A^2 \eta}{c} \theta^* (h^*)^{2\eta - 1} \tag{A.6}$$

in which (A.5) can be substituted into (A.6) to have

$$h^* = \left[\frac{\frac{A^2}{\delta c}}{\rho + \delta(1 - \eta)} \right]^{\frac{1}{2-2\eta}} > 0 \tag{A.7}$$

according to which the steady-state value h^* exists and is unique. For the other coordinate of the steady state we have

$$\theta^* = \frac{\delta c}{A^2} \left[\frac{\frac{A^2}{\delta c}}{\rho + \delta(1 - \eta)} \right]^{\frac{1-2\eta}{2-2\eta}} > 0$$

which exists and is unique as long as $\eta \neq 0.5$. With both h^* and θ^* being finite, the transversality conditions

$$\lim_{t \to \infty} e^{-\rho t} \theta(t) \geq 0 \text{ and } \lim_{t \to \infty} e^{-\rho t} \theta(t) h(t) = 0$$

obviously hold.

□

Appendix B

Note that (9) and (10) imply the equation

$$\frac{A^2 \eta}{c} h^{2\eta} \theta^2 - (\rho + \delta) h \theta + 1 = 0 = \dot{\theta} \tag{B.1}$$

which defines the roots of a quadratic polynomial in θ. In the $S = (h, \theta)$ space the $\dot{\theta} = 0$ nullcline is where we find those points that satisfy (B.1). Depending on h, equation (B.1) might have zero, one or two real roots. With simple algebra we have the roots given as

$$\frac{(\rho+\delta)h \pm \sqrt{(\rho+\delta)^2 h^2 - 4\frac{A^2\eta}{c} h^{2\eta}}}{2\frac{A^2\eta}{c} h^{2\eta}}. \tag{B.2}$$

Equations (B.1) and (B.2) define a horizontally positioned and distorted parabolic curve the vertex of which is above a stock $\tilde{h} > 0$ where the radicand is zero. This condition is easy to translate into a unique capital stock

$$\tilde{h} = \left[\frac{4\frac{A^2\eta}{c}}{(\rho+\delta)^2}\right]^{\frac{1}{2-2\eta}} > 0.$$

It is also easy to realize that the radicand is strictly increasing in h if

$$h > \left[\frac{4\frac{A^2\eta^2}{c}}{(\rho+\delta)^2}\right]^{\frac{1}{2-2\eta}}$$

in which the lower bound for any $0 < \eta < 1$ is strictly lower than $\tilde{h}$. Therefore, in our relevant range $\tilde{h} \leq h \leq \bar{h}$, the radicand in (B.2) is strictly increasing. From (B.2) we thus have the pair of equations

$$\begin{aligned}\theta_1(h) &= \frac{(\rho+\delta)h + \sqrt{(\rho+\delta)^2\, h^2 - 4\frac{A^2\eta}{c}h^{2\eta}}}{2\frac{A^2\eta}{c}h^{2\eta}} \\ \theta_2(h) &= \frac{(\rho+\delta)h - \sqrt{(\rho+\delta)^2\, h^2 - 4\frac{A^2\eta}{c}h^{2\eta}}}{2\frac{A^2\eta}{c}h^{2\eta}}\end{aligned} \tag{B.3}$$

describing the two distinct segments of the $\dot{\theta} = 0$ nullcline with $\theta_1(h)$ and $\theta_2(h)$ being the upper and the lower segment, respectively. They meet above $\tilde{h}$ where $\theta_1(\tilde{h}) = \theta_2(\tilde{h}) > 0$, and their domain is the subset $\tilde{h} \leq h \leq \bar{h}$ of the real line.

By factoring out h in the numerator of the equation for $\theta_2(h)$, we realize that by

$$(\rho+\delta)h - h\sqrt{(\rho+\delta)^2 - 4\frac{A^2\eta}{c}\frac{1}{h^{2-2\eta}}}$$

we have $\lim_{h\to\infty} \theta_2(h) = 0$ with $\theta_2(h)$ asymptotically approaching zero from above, so $\theta_2(h) > 0$ all over its domain. Differentiating $\theta_2(h)$ with respect to h yields

$$\frac{c\left[\sqrt{(\delta+\rho)^2h^2 - 4\frac{A^2\eta}{c}h^{2\eta}} - (\delta+\rho)h\right]}{A^2h^{2\eta+1}} + \frac{c\left[(\delta+\rho) - \frac{2h(\delta+\rho)^2 - \frac{8A^2\eta^2h^{2\eta-1}}{c}}{2\sqrt{(\delta+\rho)^2h^2 - 4\frac{A^2\eta}{c}h^{2\eta}}}\right]}{2A^2\eta h^{2\eta}}$$

the first part of which is negative, and the second part of which approaches $-\infty$ when $h \downarrow \tilde{h}$ as the radicand in the numerator tends to zero. It shows that over $\tilde{h}$ the slope of $\theta_2(h)$ is $-\infty$. It is also easy to demonstrate that $\theta_2(h)$ is strictly decreasing, and that $\lim_{h\to\infty} \theta_2'(h) = 0$.

Plausibly, these characteristics taken together suggest that $\theta_2(h)$ is strictly convex, but formally no general statements about convexity can be made. The characteristics of $\theta_2(h)$ over $\tilde{h}$ and at

infinity excludes the possibility of strict concavity, so we only need to exclude weak convexity by showing that the curve has neither straight segments nor kinks. Both would mean that there is more than one h where the first differential of $\theta_2(h)$ has the same value, but this problem can only be solved numerically for distinct parameter values. Weak convexity could also be excluded by showing that the second differential cannot be zero, but the most we can have is a bunch of numerical solutions to h depending on the parameters that show that the second differential cannot be zero to any $h \geq \tilde{h}$. A third strategy could be built on the first-order Taylor-polynomial underestimating $\theta_2(h)$ at any h if it was strictly convex. Along this path, we cannot show uniqueness of any h where $\theta_2(h)$ and the Taylor polynomial have osculation.

Monotonicity of $\theta_1(h)$ crucially depends on η. It is obvious that for $0 < \eta < 0.5$ the function $\theta_1(h)$ in (B.3) is the sum of two strictly increasing terms, so it is strictly increasing. By simple algebra (B.3) can also be rewritten as

$$\theta_1(h) = \frac{(\rho + \delta) + \sqrt{(\rho + \delta)^2 - 4\frac{A^2\eta}{c}h^{2\eta-2}}}{2\frac{A^2\eta}{c}h^{2\eta-1}}$$

in which the numerator tends to $2(\rho + \delta)$ when $h \to \infty$ for any $0 < \eta < 1$. The term in the denominator, however, tends to infinity if $1 > \eta > 0.5$ hence $\theta_1(h) \to 0$ in this case. By contrast, if $0 < \eta < 0.5$, the asymptotic behavior of $\theta_1(h)$ is given by

$$\lim_{h\to\infty} \theta_1(h) = \frac{c(\rho+\delta)}{A^2\eta}h^{1-2\eta}. \tag{B.4}$$

Therefore, for $0 < \eta < 0.5$ the curve of $\theta_1(h)$ is strictly increasing tending towards infinity when $h \to \infty$, but asymptotically approaches zero in cases where $0.5 < \eta < 1$. The first-order differential of $\theta_1(h)$ clearly shows that its slope over $\tilde{h}$ is $+\infty$, so in the $0.5 < \eta < 1$ case there must be a finite h where $\theta_1(h)$ has a global maximum. By direct calculation we can show that this unique value of $h > 0$ is

$$\left[\frac{c(2\eta-1)^2(\delta+\rho)^2}{4A^2\eta^3}\left(\frac{2\eta}{2\eta-1}-1\right)\right]^{\frac{1}{2\eta-2}}.$$

The uniqueness of the steady state that we demonstrated in Proposition 1 above ensures that even if all three curves tend to zero at infinitely high values of h, there is only one point of intersection, so the two sections of the $\dot{\theta} = 0$ nullcline keeps the $\dot{h} = 0$ nullcline trapped in the $0.5 < \eta < 1$ case. It can also be shown that the point (h^*, θ^*) is always on $\theta_2(h)$.

Appendix C

Proposition 2: *For* $0 < \eta < 0.5$*, the steady state characterized by the point* (h^*, θ^*) *in the* $S = (h, \theta)$ *space is saddle-path stable, so for any* h_0 *such that* $1 < h_0 < h^*$ *there can be found a unique* $\theta(0)$

from which the system defined by (5) and (6) converges to (h^*, θ^*) *along a stable manifold with* $\dot{h} > 0$ *and* $\dot{\theta} < 0$. *The optimal path toward the steady state is unique.*

Proof: First, we need to understand what happens to $\dot{h}$ above and below the $\dot{h} = 0$ nullcline. As for any $0 < \eta < 0.5$ the $\dot{h} = 0$ nullcline takes all values of θ between 0 and $+\infty$, the nullcline halves the positive quadrant of the $S = (h, \theta)$ space. By using this fact, and the strict monotonicity of the nullcline, for any h we can find a unique θ identifying a unique point in S where $\dot{h} = 0$ and where we can leave the nullcline vertically. Therefore, at any h, we can differentiate (A.1) with respect to θ that yields

$$\dot{h}_\theta = \frac{A^2}{c} h^{2\eta} > 0$$

implying that $\dot{h} > 0$ above the $\dot{h} = 0$ nullcline.

Second, we need to understand what happens to $\dot{\theta}$ to the left and to the right of the $\dot{\theta} = 0$ nullcline. By the same token as in the first part of the proof, at any θ we can find an h such that $\dot{\theta} = 0$. To understand what happens to $\dot{\theta}$ when we leave the nullcline to the left, we must differentiate (A.3) with respect to h that yields

$$\dot{\theta}_h = -\frac{A^2 \eta}{c} (2\eta - 1)\theta^2 h^{2\eta - 2} + h^{-2} > 0$$

for any $0 < \eta < 0.5$. It means that if we horizontally leave the $\dot{\theta} = 0$ nullcline to the left at any given θ by reducing h, then $\dot{\theta}$ also reduces, and $\dot{\theta} < 0$ emerges. The system thus turns out to be saddle-path stable with clear directions of flows in the phase diagram. Dynamics of the system is effectively summarized in Fig. 10 below.

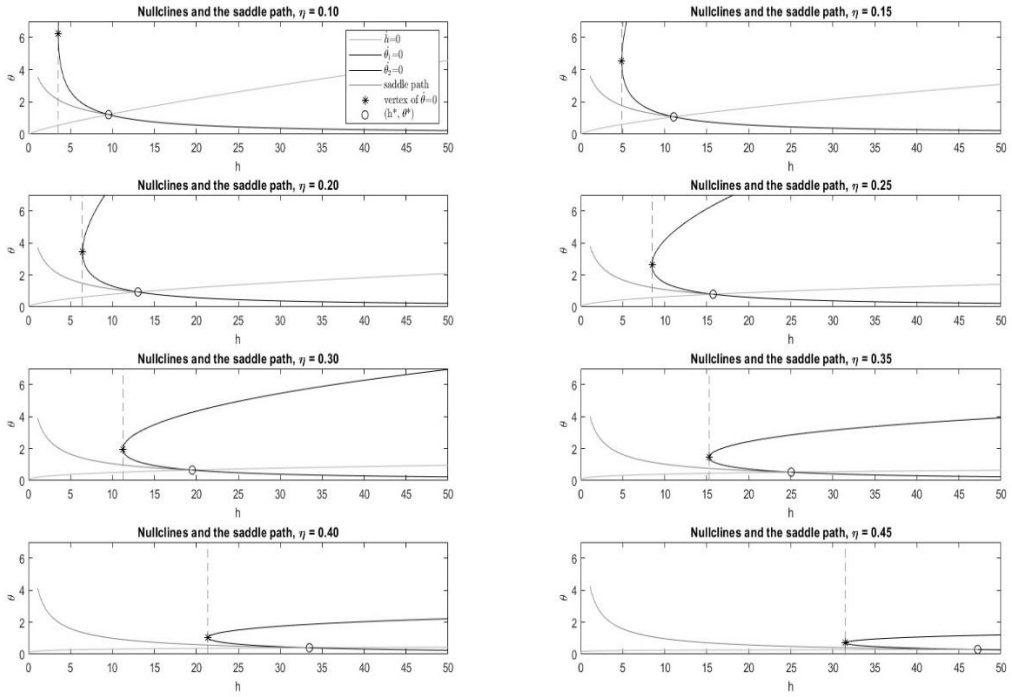


Fig. 10: Saddle-path stability in the model of school learning.

From this analysis, $\dot{h} > 0$ and $\dot{\theta} < 0$ for any $h < h^*$ emerge as the key result which is our starting point in the analysis of the optimum path of capital. For the other case, where $0.5 < \eta < 1$, we do not have clear results. It is easy to find such a point along the $\dot{\theta} = 0$ nullcline, where leaving the nullcline to the left we find ourselves to the right of the same nullcline (see point A in Fig. 2, Pane B above); or, likewise, a point, where leaving the nullcline upwards we find ourselves under the same nullcline (B). We thus disregard the $0.5 < \eta < 1$ case.

The linearization of the implied dynamics reassures saddle-path stability in the $0 < \eta < 0.5$ case. The parameter matrix of the linearized system around the steady state is the Jacobian

$$J = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} = \begin{bmatrix} \dot{h}_h & \dot{h}_\theta \\ \dot{\theta}_h & \dot{\theta}_\theta \end{bmatrix} = \begin{bmatrix} \dfrac{2\eta A^2}{c}\theta^*(h^*)^{2\eta-1} - \delta & \dfrac{A^2}{c}(h^*)^{2\eta} \\ (1-2\eta)\dfrac{\eta A^2}{c}(\theta^*)^2(h^*)^{2\eta-2} + (h^*)^{-2} & (\rho+\delta) - \dfrac{\eta A^2}{c}2\theta^*(h^*)^{2\eta-1} \end{bmatrix}$$

in which $\dot{h}_h$, $\dot{h}_\theta$, etc. denote the partial first-order differentials of (11) and (10) evaluated in the steady state. Obviously, $J_{12} > 0$, and $J_{21} > 0$ also applies for $0 < \eta < 0.5$. Substituting h^* and θ^* into J_{11} and J_{22} by direct calculation we receive that $J_{11} = \delta(2\eta - 1) < 0$ and $J_{22} = \rho + \delta(1 - 2\eta) > 0$ for the relevant case $0 < \eta < 0.5$, so the general form of the Jacobian is

$$J = \begin{bmatrix} - & + \\ + & + \end{bmatrix}$$

from which it follows that the determinant of J is negative. Using the basic fact that the determinant of J is equal to the product of its eigenvalues, it follows that one eigenvalue is strictly positive while the other is strictly negative. Moreover, the negative determinant also implies that both eigenvalues are real (Horn & Johnson, 2013, pp. 50-51). Consequently, with two real eigenvalues having opposite signs, the system is saddle-path stable in a sufficiently large neighborhood of the steady state. It is also straightforward to prove that the elements in the eigenvector appertaining to the stable eigenvalue $\lambda_s < 0$ have mixed signs. The eigenvector $[v_1\ v_2]^T$ solves the matrix equation

$$\begin{bmatrix} \lambda_s < 0 & 0 \\ 0 & \lambda_s < 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} J_{11} < 0 & J_{12} > 0 \\ J_{21} > 0 & J_{22} > 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

implying two linear equations. We want to prove that v_1 and v_2 have opposite signs. To play on a contradiction by assuming that their signs are the same, i.e. $v_i > 0$ or $v_i < 0$, $i = 1,2$ (all eigenvectors are real). In the second equation above, if $v_i > 0$, then $J_{21}v_1 + J_{22}v_2 > 0$, while $\lambda_s v_2 < 0$, which is a contradiction. In the opposite case, if $v_i < 0$, then $J_{21}v_1 + J_{22}v_2 < 0$, while $\lambda_s v_2 > 0$, which is a contradiction, so v_1 and v_2 must have opposite signs. It implies that when initially the state deviates from its steady-state value in the negative direction along the stable manifold, i.e. $\tilde{h}(0) = h_0 - h^* < 0$, then the costate deviates in the positive direction, i.e. $\tilde{\theta}(0) = \theta(0) - \theta^* > 0$. Formally put, given that $\tilde{h}(t) = \tilde{h}(0)e^{\lambda_s t}$, we have $\dot{\tilde{h}}(t) = \tilde{h}(0)\lambda_s e^{\lambda_s t} > 0$, while $\dot{\tilde{\theta}}(t) = \frac{v_2}{v_1}\tilde{h}(0)\lambda_s e^{\lambda_s t} < 0$ following from $\tilde{\theta}(t) = \frac{v_2}{v_1}\tilde{h}(t) = \frac{v_2}{v_1}\tilde{h}(0)e^{\lambda_s t}$ where $\tilde{h}(0) < 0$, $\lambda_s < 0$ and $v_2/v_1 < 0$. Therefore, along the stable manifold $\dot{h}(t) > 0$ starting from $h_0 < h^*$ and $\dot{\theta}(t) < 0$.

As a final step in the proof, when $0 < \eta < 0.5$, the optimized Hamiltonian is strictly concave in h, so the optimal path is unique. To see this, let $\hat{u}$ denote the path of the control variable that maximizes the Hamiltonian in (8). Then, the maximized Hamiltonian is

$$H^0(h;\theta;u) \equiv H(h;\theta;\hat{u}) = \ln h - \frac{c}{2}\hat{u}^2 + \theta(Ah^\eta \hat{u} - \delta h)$$

the second differential of which with respect to h is $-h^{-2} + \eta(\eta - 1)A\theta h^{\eta-2}\hat{u} < 0$ that implies the strict concavity of H^0. The optimal path towards the steady state is thus unique.

□

Appendix D

By using (6) and (9), the second differential of h is

$$\ddot{h} = \frac{A^2}{c}\left[\dot{\theta}h^{2\eta} + 2\eta\theta h^{2\eta-1}\dot{h}\right] - \delta\dot{h} \qquad \text{(D.1)}$$

which might be positive hence the optimum path of h might be strictly concave or might have convex sections as $\dot{\theta} < 0$ and $\dot{h} > 0$ all along the path with $\lim_{t\to\infty} \dot{h}(t) = 0$ and $\lim_{t\to\infty} \dot{\theta}(t) = 0$.

Substituting from (10) and (11) we can trace (D.1) back to the dynamics of h and θ, and receive an expression that can be used as a criterion for convexity, given as

$$\frac{A^2\eta}{c}(\theta h^{2\eta-1})^2 + (\rho - 2\eta\delta)\theta h^{2\eta-1} + \frac{c\delta^2}{A^2} > h^{2\eta-2} \qquad \text{(D.2)}$$

in which the inequality holds for convex sections of the path of h, and the opposite holds for concave sections. Note that the left-hand side of (D.2) is a second-order polynomial of the variable $\theta h^{2\eta-1}$, the roots of which are determined by

$$\frac{-(\rho - 2\eta\delta) \pm \sqrt{(\rho - 2\eta\delta)^2 - 4\frac{A^2\eta}{c}\frac{c\delta^2}{A^2}}}{\frac{2A^2\eta}{c}}$$

in which $(\rho - 2\eta\delta)$ is likely to be positive for sufficiently low values of δ, so the polynomial has no positive roots. A strictly concave path of h is summarized in Fig. 11 below:

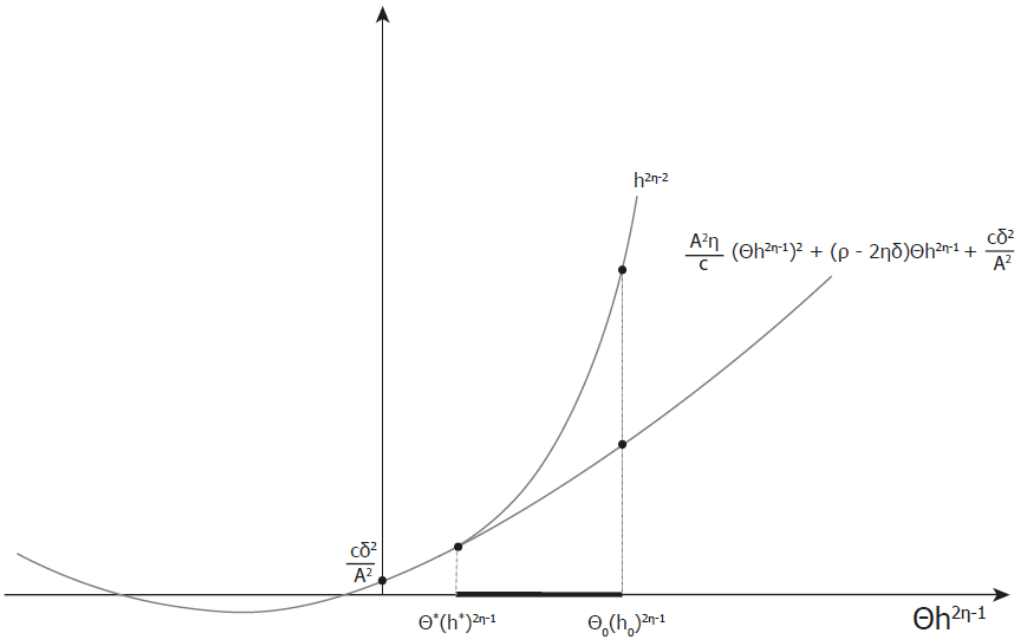


Fig. 11: If the optimum path of h is strictly concave, then the second-order polynomial in $\theta h^{2\eta-1}$ remains under the curve $h^{2\eta-2}$.

As $\dot{\theta} < 0$ and $\dot{h} > 0$, and $0 < \eta < 0.5$, the ratio $\theta h^{2\eta-1}$ (or $\theta/h^{1-2\eta}$) has its highest value at h_0 and θ_0, and the learning process reduces it to $\theta^*/(h^*)^{1-2\eta}$, which by the use of (A.7) turns out to be $\delta c/A^2$ in all possible cases. Therefore, the close end of the state space defined in $\theta/h^{1-2\eta}$ is fixed in this point. Moreover, any point of an optimum path can be represented as a unique point in $\theta/h^{1-2\eta}$ since by the strict monotonicity of the dynamics of the state and the costate, any pairs of h and θ takes a unique value in $\theta/h^{1-2\eta}$. From this it follows that along any optimum path the

position of $h^{2\eta-2}$ in Fig. 11 is clearly defined since to any point in the form $\theta/h^{1-2\eta}$ there belong a unique h, so for strict concavity $h^{2\eta-2}$ works as an upper bound for the left-hand side of (D.2). By direct calculation, (D.2) turns out to hold with equality over $\theta^*/(h^*)^{1-2\eta}$ (i.e. at the close end), so here $\ddot{h}=0$ hence at the end of the optimum path it is neither concave nor convex. It is also straightforward to point out that $\dot{h}=0$ in h^*, and these two features together with $\dot{h}>0$ under the optimum capital stock imply that strict convexity of $h(t)$ can be excluded. Right before reaching h^* the very latest a flat section of $h(t)$ inevitably emerges. If there are convex sections, they occur at the earlier phases of human capital accumulation. They possibly exist as by differentiating (6) with respect to time at $u=1$ we receive a strictly convex section reaching up to relatively high values of h. The easiest way to find a convex section is to elevate the starting point of the left-hand side expression in (D.2) over $h^{2\eta-2}$, if possible. By doing so, we could find an optimum path which starts with a convex section fading away as in the optimum capital stock, as we have just noted, the left-hand side expression approaches $h^{2\eta-2}$ from below. Multiple convex sections cannot be excluded. The most we can know about a partly convex case is the possible existence of an initial convex phase and a final flat section along the optimum path.

By changing the starting values h_0 and $\theta(0)$ we might therefore find a far end for the state space in $\theta/h^{1-2\eta}$ in Fig. 11 that implies (D.2) with inequality signaling convexity. To see this, two details must be noted. First, following from $\dot{\theta}<0$, a lower $h_0>1$ implies a higher $\theta(0)>0$ *ceteris paribus* as a starting value along a stable manifold. And second, even if two students start from the same h_0, a higher η ensures possession of a higher h^*, and the higher difference between h_0 and the optimum knowledge levels implies that a higher η makes the same h_0 relatively scarcer and hence more valuable, leading to a higher $\theta(0)$.

Appendix E

There are parameter configurations under which the untenable exterior solution $u(t)>1$ arises for some t. In such cases, the upper bound $u(t)=1$ must be active in the system of differential equations described in (10) and (11). To have meaningful solutions, however, we must check if the constrained process tends toward a unique and stable solution h_c^* and θ_c^*, where c in the lower index signals the constrained process.

Proposition 3: *The constrained version of the problem defined by (5) and (6) where* $u(t)=1$ *has a unique steady state* h_c^* *and* θ_c^**. These solutions are finite at any configuration of appropriate and finite parameters.*

Proof: In Section 3, we showed that there is a unique and finite maximum maintainable capital stock $\bar{h}$ that we received by setting $u(t)=1$ and $\dot{h}=0$ in (6) and (11). Therefore, now we just need to add that $h_c^*\equiv\bar{h}$ which, at the same time, is the $\dot{h}_c=0$ nullcline of the constrained process. When $u(t)=1$ is applied to (10), the dynamics of the costate is described by the equation

$$\dot{\theta}_c = \theta\left(\rho + \delta - A\eta h_c^{\eta-1}\right) - \frac{1}{h_c}. \tag{E.1}$$

If there is a steady state, it is on both the $\dot{h}_c = 0$ and $\dot{\theta}_c = 0$ nullclines, so $h_c = h_c^*$, $\theta_c = \theta_c^*$, and $\dot{\theta}_c = 0$ in (E.1). When $\dot{\theta}_c = 0$ while h_c and θ_c are free in (E.1), we have those h_c and θ_c pairs where $\dot{\theta}_c = 0$, hence we have the $\dot{\theta}_c = 0$ nullcline. From this, it is straightforward to show that

$$\theta_c^* = \frac{\left(\frac{\delta}{A}\right)^{\frac{1}{1-\eta}}}{\rho + \delta(1-\eta)} = \frac{h_c^*}{\rho + \delta(1-\eta)}$$

which is unique and finite for h_c^* being finite and $\rho + \delta(1-\eta) > 0$.

Now that the system in (10) and (11) certainly converges toward a unique steady state with the $u(t) = 1$ constraint, the stability of the steady state is to be checked. Just like in the case of the unconstrained problem, we can proceed by linearizing (E.1) and

$$\dot{h}_c = Ah^\eta - \delta h. \tag{E.2}$$

around the steady state.

Proposition 4: *The steady state h_c^* and θ_c^* of the constrained process is stable.*

Proof: By differentiating (E.1) and (E.2) with respect to h and θ, the constrained process turns out to have the Jacobian

$$J^c = \begin{bmatrix} J_{11}^c & J_{12}^c \\ J_{21}^c & J_{22}^c \end{bmatrix} = \begin{bmatrix} (\dot{h}_c)_h & (\dot{h}_c)_\theta \\ (\dot{\theta}_c)_h & (\dot{\theta}_c)_\theta \end{bmatrix} = \begin{bmatrix} \delta(\eta-1) & 0 \\ A\eta(1-\eta)\dfrac{\left(\frac{A}{\delta}\right)^{\frac{\eta-3}{1-\eta}}}{\rho + \delta(1-\eta)} + \left(\frac{A}{\delta}\right)^{\frac{-2}{1-\eta}} & \rho + \delta(1-\eta) \end{bmatrix} = \begin{bmatrix} - & 0 \\ + & + \end{bmatrix}$$

whose determinant is negative, it thus has two real and nonzero eigenvalues with opposite signs, so the constrained process is saddle path stable—to any starting value $1 < h_c(0) < h_c^*$ (as the analogue of h_0 in the unconstrained problem) there belongs a unique $\theta_c(0)$ from which the paths for $h_c(t)$ and $\theta_c(t)$ start to converge to h_c^* and θ_c^*, respectively. The dynamics of $\dot{h}_c$ in (E.2) does not depend on $\dot{\theta}_c$, but $\dot{h}_c$ enters the law of motion of $\dot{\theta}_c$, and accordingly under any appropriate parameter configuration specified earlier there is a unique path for $\theta_c(t)$ along which the constrained process reaches the steady state. If the stable eigenvalue is λ_s^c, then the eigenvector $[v_1^c \; v_2^c]^T$ solves the equation

$$\begin{bmatrix} \lambda_s^c < 0 & 0 \\ 0 & \lambda_s^c < 0 \end{bmatrix} \begin{bmatrix} v_1^c \\ v_2^c \end{bmatrix} = \begin{bmatrix} J_{11}^c < 0 & J_{12}^c = 0 \\ J_{21}^c > 0 & J_{22}^c > 0 \end{bmatrix} \begin{bmatrix} v_1^c \\ v_2^c \end{bmatrix}.$$

To show that $\dot{h}_c(t) > 0$ and $\dot{\theta}_c(t) < 0$, we can apply the results in the proof of Proposition 2 by showing that v_1^c and v_2^c have opposite signs. By assuming that the signs are the same, we get a contradiction in either case, so v_1^c and v_2^c have opposite signs, hence $\dot{h}_c(t) > 0$ and $\dot{\theta}_c(t) < 0$.

In the simplest case, $\dot{u}(t) < 0$ in the unconstrained process with possibly $u(t) > 1$ for a period $t \in [0, T)$ with the condition $u(T) = 1$, so T is the time where the exterior path of $u(t)$ becomes interior. To this, due to the saddle path stability of the unconstrained problem, there belongs a unique pair of $h(T)$ and $\theta(T)$. In such a case, we can halve the original unconstrained problem. For the first period with exterior solution, we resort to the constrained problem to describe the dynamics of human capital with $\dot{h}_c(t) > 0$ from the starting value of the unconstrained problem, hence $h_c(0) = h_0$. As $\dot{h}_c(t) > 0$, the constrained dynamics once reaches $h(T)$ where the unconstrained part of the problem starts with $u(t) \leq 1$ all along until h^* and θ^* are reached. Therefore, $h(T)$ of the original problem with exterior solution becomes the boundary condition for the unconstrained part of the problem. We switch from the constrained to the unconstrained problem with a jump in the costate at T. This pattern can be generalized to any case.

Appendix F

Proposition 5: *There exists a unique maximum level of knowledge* $\bar{h}_i$, $i = 1,2$, *to each agent over which no capital stock is maintainable to do the constraint* $0 \leq u_i(t) \leq 1$.

Proof: These upper bounds are where $\dot{h}_1 = \dot{h}_2 = 0$ and $u_1 = u_2 = 1$. It follows that we have a two-equation system for the maximum levels

$$A\bar{h}_2^{\frac{\gamma}{2}}\bar{h}_1^{\eta_1+\frac{\gamma}{2}} = \delta\bar{h}_1 \tag{F.1a}$$

$$A\bar{h}_1^{\frac{\gamma}{2}}\bar{h}_2^{\eta_1+\frac{\gamma}{2}} = \delta\bar{h}_2 \tag{F.1b}$$

that is to be solved for maximum levels $\bar{h}_1$ and $\bar{h}_2$. From (F.1a) it follows that

$$\bar{h}_1 = \left(\frac{A}{\delta}\right)^{\frac{1}{1-\eta_1-\frac{\gamma}{2}}} \bar{h}_2^{\frac{\gamma}{2\left(1-\eta_1-\frac{\gamma}{2}\right)}} \tag{F.2}$$

hence $\bar{h}_1 = f(\bar{h}_2)$, where

$$0 < \frac{\gamma}{2\left(1-\eta_1-\frac{\gamma}{2}\right)} < 1,$$

so $\bar{h}_1$ is an increasing function of the potentially infinite $\bar{h}_2$ (note that we should not presume the finiteness of any $\bar{h}_i$), hence it is also potentially infinite. Using symmetry, we also have

$$\bar{h}_2 = \left(\frac{A}{\delta}\right)^{\frac{1}{1-\eta_2-\frac{\gamma}{2}}} \bar{h}_1^{\frac{\gamma}{2\left(1-\eta_2-\frac{\gamma}{2}\right)}} \tag{F.3}$$

hence $\bar{h}_2 = g(\bar{h}_1)$, and eventually $\bar{h}_1 = f\left(g(\bar{h}_1)\right)$, so the finite upper bound, if it exists, emerges as a fixed point of the mapping $f(g(\cdot))$. It is straightforward to combine (F.2) and (F.3) to have

$$\bar{h}_1 = \left(\frac{A}{\delta}\right)^{\frac{1}{1-\eta_1-\frac{\gamma}{2}}} \left(\left(\frac{A}{\delta}\right)^{\frac{1}{1-\eta_2-\frac{\gamma}{2}}} \bar{h}_1^{\frac{\gamma}{2\left(1-\eta_2-\frac{\gamma}{2}\right)}}\right)^{\frac{\gamma}{2\left(1-\eta_1-\frac{\gamma}{2}\right)}} = \left(\frac{A}{\delta}\right)^{\frac{1-\eta_2}{\left(1-\eta_1-\frac{\gamma}{2}\right)\left(1-\eta_2-\frac{\gamma}{2}\right)}} \bar{h}_1^{\frac{\gamma^2}{4\left(1-\eta_1-\frac{\gamma}{2}\right)\left(1-\eta_2-\frac{\gamma}{2}\right)}}$$

which is the explicit formulation of the fixed-point problem $\bar{h}_1 = f\left(g(\bar{h}_1)\right)$. By simple manipulation it becomes obvious that

$$\bar{h}_1 = \left(\frac{A}{\delta}\right)^{\frac{4-4\eta_2}{4\left(1-\eta_1-\frac{\gamma}{2}\right)\left(1-\eta_2-\frac{\gamma}{2}\right)-\gamma^2}},$$

and by the symmetry of the solution, it also follows that

$$\bar{h}_2 = \left(\frac{A}{\delta}\right)^{\frac{4-4\eta_1}{4\left(1-\eta_1-\frac{\gamma}{2}\right)\left(1-\eta_2-\frac{\gamma}{2}\right)-\gamma^2}}.$$

Both $\bar{h}_1$ and $\bar{h}_2$ are finite and unique, so we can expect to have finite solutions to the equilibrium knowledge levels.

□

Appendix G

Proposition 6: *Regarding the equilibrium capital stock* h_1^e*, the differential* $(h_1^e)_{\eta_2} > 0$ *applies as long as the condition*

$$\frac{\frac{A^2}{\delta c}}{\rho + \delta\left(1-\eta_i-\frac{\gamma}{2}\right)} > 1 \tag{G.1}$$

holds where $i = 1,2$.

Proof: Note that η_2 shows up at three places in (20). Instead of calculating $(h_1^e)_{\eta_2}$ directly, to simplify the problem we might first try to calculate the differentials of these parts. Accordingly, we have

$$\left[\frac{\frac{A^2}{\delta c}}{\rho + \delta\left(1-\eta_2-\frac{\gamma}{2}\right)}\right]_{\eta_2}' = \frac{\frac{A^2}{c}}{\left[\rho + \delta\left(1-\eta_2-\frac{\gamma}{2}\right)\right]^2} > 0 \tag{G.2}$$

$$\left[\frac{2-2\eta_2-\gamma}{(2\eta_1+\gamma-2)(2\eta_2+\gamma-2)-\gamma^2}\right]'_{\eta_2} = \frac{\gamma^2}{2[\gamma\eta_1-2\gamma-2\eta_1+\eta_2(2\eta_1+\gamma-2)+2]^2} > 0 \tag{G.3}$$

$$\left[\frac{\gamma}{(2\eta_1+\gamma-2)(2\eta_2+\gamma-2)-\gamma^2}\right]'_{\eta_2} = \frac{\gamma(4-4\eta_1-2\gamma)}{[(2\eta_1+\gamma-2)(2\eta_2+\gamma-2)-\gamma^2]^2} > 0 \tag{G.4}$$

in which the third inequality comes from the requirement that $0.5 > \eta_i + \frac{\gamma}{2}$ hence $2 > 4\eta_1 + 2\gamma$. The same requirement also implies that the denominators in (G.3) and (G.4) are strictly positive. Practically, it means that all the three terms, the second base and both exponents, are increasing functions of η_2. The basic monotonicity properties of the power function imply that h_1^e is an increasing function of η_2 as long as the two bases in (20) are above 1 according to (G.1).

□

Note that we expect (G.1) to hold following from (A.7) and the initial condition that $h_i(0) > 1$. Consequently, $(h_1^e)_{\eta_2} > 0$, so the equilibrium knowledge level of the family having superior learning abilities is an increasing function of the learning abilities of the other family in the same environment. At the intuitive level, it aligns with our intuition that a classmate with inferior learning potential holds her classmate back in human capital accumulation in an environment where learning abilities interplay through an external effect.

We can use similar arguments to show $(h_1^e)_{\eta_1} > 0$.

Proposition 7: *Regarding the equilibrium capital stock* h_1^e*, the differential* $(h_1^e)_{\eta_1} > 0$ *holds as long as the condition in (G.1) holds where* $i = 1,2$.

Proof: Due to the symmetry in (20) in terms of η_1 and η_2, the direct analogues of (G.2) with η_1 replacing η_2 and (G.4) being indifferent if differentiated with respect to η_1 or η_2 directly apply. For the third term involved, we have

$$\left[\frac{2-2\eta_2-\gamma}{(2\eta_1+\gamma-2)(2\eta_2+\gamma-2)-\gamma^2}\right]'_{\eta_1} = \frac{2(2\eta_2+\gamma-2)^2}{[(2\eta_1+\gamma-2)(2\eta_2+\gamma-2)-\gamma^2]^2} > 0$$

in which the denominator is strictly positive due to the same reason as we used in the case of (G.4). As long as (G.1) holds, the monotonicity of the power function implies that $(h_1^e)_{\eta_1} > 0$.

□

The symmetry in (20a) and (20b) allows us to claim that $(h_2^e)_{\eta_1} > 0$ as the analogue of Proposition 6, and that $(h_2^e)_{\eta_2} > 0$ as the analogue of Proposition 7. From (20a) and (20b) it is also obvious that the assumption $\eta_1 > \eta_2$ does not affect the generality of the results. As $(h_1^e)_{\eta_2} > 0$ and $(h_2^e)_{\eta_1} > 0$ while $\eta_1 > \eta_2$ holds, we can easily relax the assumption by replacing η_1 and η_2 with each other in (20a) and (20b). Accordingly, if Agent 1 is exposed to one, say, an Agent 2 having

higher learning abilities, Agent 1 takes over the role of Agent 2 in (20b), but rewriting (20b) with η_1 replacing η_2 and *vice versa*, we receive (20a), but now with $\eta_2 > \eta_1$. Consequently, all we have to say is that the equilibrium stock of each Agent i is positively related to η_j of Agent j, $i = 1,2$ and $j = 1,2$, where $i \neq j$.

Appendix H

Proposition 8: *The equilibrium allocation with* h_1^e, h_2^e, θ_1^e *and* θ_2^e *is stable.*

Proof: For the formal proof, suffice it to turn back to the linearized dynamics of the isolated individual as carried out in Appendix C. In the equilibrium setting, we have an analogue Jacobian with which we check if, moved out of the equilibrium, the individual's capital stock returns to the equilibrium state. Given (19a) and (19b), in place of J_{11} we have

$$\frac{2\left(\eta_1+\frac{\gamma}{2}\right)A^2(h_2^e)^\gamma}{c}\frac{\delta c}{A^2(h_2^e)^\gamma}\left[\frac{\frac{A^2(h_2^e)^\gamma}{\delta c}}{\rho+\delta\left(1-\eta_1-\frac{\gamma}{2}\right)}\right]^{\frac{1-2\left(\eta_1+\frac{\gamma}{2}\right)}{2-2\left(\eta_1+\frac{\gamma}{2}\right)}}\left[\frac{\frac{A^2(h_2^e)^\gamma}{\delta c}}{\rho+\delta\left(1-\eta_1-\frac{\gamma}{2}\right)}\right]^{\frac{2\left(\eta_1+\frac{\gamma}{2}\right)-1}{2-2\left(\eta_1+\frac{\gamma}{2}\right)}}-\delta$$
$$=2\left(\eta_1+\frac{\gamma}{2}\right)\delta-\delta=\delta\left[2\left(\eta_1+\frac{\gamma}{2}\right)-1\right]<0$$

for agent 1 with $0 < \eta_1 + \gamma/2 < 0.5$ by assumption. By the symmetry of the problem, for Agent 2 we of course have $\delta[2(\eta_2 + \gamma/2) - 1] < 0$ in place of J_{11}. In place of J_{22} we have

$$(\rho+\delta)-\frac{\left(\eta_1+\frac{\gamma}{2}\right)A^2(h_2^e)^\gamma}{c}2\frac{\delta c}{A^2(h_2^e)^\gamma}\left[\frac{\frac{A^2(h_2^e)^\gamma}{\delta c}}{\rho+\delta\left(1-\eta_1-\frac{\gamma}{2}\right)}\right]^{\frac{1-2\left(\eta_1+\frac{\gamma}{2}\right)}{2-2\left(\eta_1+\frac{\gamma}{2}\right)}}\left[\frac{\frac{A^2(h_2^e)^\gamma}{\delta c}}{\rho+\delta\left(1-\eta_1-\frac{\gamma}{2}\right)}\right]^{\frac{2\left(\eta_1+\frac{\gamma}{2}\right)-1}{2-2\left(\eta_1+\frac{\gamma}{2}\right)}}$$
$$=(\rho+\delta)-2\delta\left(\eta_1+\frac{\gamma}{2}\right)=\rho+\delta\left[1-2\left(\eta_1+\frac{\gamma}{2}\right)\right]>0$$

for agent 1 with $0 < \eta_1 + \gamma/2 < 0.5$ by the same assumption. For Agent 2, analogously, we have $\rho + \delta[1 - 2(\eta_2 + \gamma/2)] > 0$ in place of J_{11}. It is straightforward to see that the social equilibrium analogues of J_{12} and J_{21} are both positive, so the equilibrium in the social setting is saddle path stable for both agents.

Appendix E

Proposition 9: *Function* $h(h_1^*)$ *in (29) uniquely determines the optimal capital stock* $h_1^* > 0$.

Proof: Suffice it to show that $h(h_1^*)$ is strictly increasing in h_1^*. If so, there is a unique h_1^* where the curve of $h(h_1^*)$ crosses the horizontal line on the left. The positivity of the multiplier in the braces in (29) is straightforward to prove by assuming positivity and demonstrating that no contradiction follows. This assumption directly implies that

$$\frac{2\left[\delta\rho c+\delta^2 c\left(1-\eta_1-\frac{\gamma}{2}\right)\right]}{\delta^2\gamma c} > \frac{2\left[\delta\rho c+\delta^2 c\left(1-\eta_1-\frac{\gamma}{2}\right)\right]+\delta^2\gamma c}{2\left[\delta\rho c+\delta^2 c\left(1-\eta_2-\frac{\gamma}{2}\right)\right]+\delta^2\gamma c}.$$

After expansion and simplification, we can realize that all terms in the inequality are $c^2\delta^2$, $c^2\delta^3$ and $c^2\delta^4$ types. By grouping the expressions, it turns out that the multiplier in (29) is positive if

$$2c^2\delta^4[2(1-\eta_1-\eta_2-\gamma)+\gamma(\eta_1+\eta_2)+2\eta_1\eta_2] > 0$$

and

$$4\delta^3\rho c^2(2-\eta_1-\eta_2-\gamma) > 0$$

of which the second inequality obviously holds due to the parameter restrictions, and the first also holds by the previous assumption that $0 < \eta_j + \frac{\gamma}{2} < 0.5$ implying $1-\eta_1-\eta_2-\gamma > 0$. To have a strictly increasing function in (29) we also need the exponent to be positive that holds as long as $(2-\gamma)(1-\eta_1-\eta_2)-\gamma+2\eta_1\eta_2 > 0$. By simple algebra, the requirement boils down to $1-\eta_1-\eta_2-\gamma > 0$. The existence and the uniqueness of $h_1^* > 0$ is proven.

□

Appendix I

Proposition 10: *The unique optimum capital stock* h_1^* *as determined by* $h(h_1^*)$ *in (29) is an increasing function of* η_2, therefore $(h_1^*)_{\eta_2} > 0$.

Proof 1: We use equation (25a) to find the differential of h_1^* with respect to η_2. In the steady state, $\dot{\theta}_1^* = 0$, therefore (25a) becomes

$$(\rho+\delta)\dot{\theta}_1^* = \frac{1}{h_1^*} + A\left(\eta_1+\frac{\gamma}{2}\right)\theta_1^*(h_1^*)^{\eta_1+\frac{\gamma}{2}-1}(h_2^*)^{\frac{\gamma}{2}}u_1^* + A\frac{\gamma}{2}\theta_2^*(h_1^*)^{\frac{\gamma}{2}-1}(h_2^*)^{\eta_2+\frac{\gamma}{2}}u_2^*$$

in which we assume that h_1^* depends on η_1. Proceeding on the basis of the envelope theorem and the differentiation of implicit functions, we disregard all the indirect effects of a change in η_1 on all other variables and get

$$\begin{aligned}(-1)(h_1^*)^{-2}(h_1^*)_{\eta_2} - A\left(1-\eta_1-\frac{\gamma}{2}\right)\left(\eta_1+\frac{\gamma}{2}\right)\theta_1^*u_1^*(h_1^*)^{\eta_1+\frac{\gamma}{2}-2}(h_2^*)^{\frac{\gamma}{2}}(h_1^*)_{\eta_2} \\ - A\left(1-\frac{\gamma}{2}\right)\frac{\gamma}{2}\theta_2^*u_2^*(h_1^*)^{\frac{\gamma}{2}-2}(h_2^*)^{\eta_2+\gamma}(h_1^*)_{\eta_2} + A\frac{\gamma}{2}\theta_2^*u_2^*(h_1^*)^{\frac{\gamma}{2}-1}(h_2^*)^{\eta_2+\frac{\gamma}{2}}\ln(h_2^*) = 0\end{aligned}$$

which is the sum of four terms, including three parameters on $(h_1^*)_{\eta_2}$. Note that under the stated assumptions $\gamma > 0$ and $0.5 > \eta_i + \frac{\gamma}{2} > 0$, and given the results that $h_i^* > 0$, $\theta_i^* > 0$ and $u_i^* > 0$, where $i = 1,2$, all the three parameters on $(h_1^*)_{\eta_2}$ are negative, while the fourth is positive, so after rearranging the equation we receive $(h_1^*)_{\eta_2} > 0$.

□

Proof 2: Note that in (29), h_1^* is described by a power function, so a sufficient condition for h_1^* to be an increasing function of η_2 in the case where $(2-\gamma)(1-\eta_1-\eta_2)-\gamma+2\eta_1\eta_2 > 0$ (as

demonstrated in Proposition 9) is that both the exponent of h_1^* and the multiplier are *decreasing* functions of η_2. In this case, the curve implied by the right-hand side of (29) flattens out and crosses the line on the left at higher values of h_1^*. This line of argument works only for cases where $h_1^* > 1$, but we can easily ensure it holds as the baseline case by the appropriate choice of parameters to have meaningful solutions for convergence with $h_1(0) > 1$.

As far as the exponent of h_1^* is concerned, by direct differentiation it is easy to show that it is a decreasing function of η_2. To see that the multiplier is also a decreasing function of η_2, we should rewrite the multiplier in (29) in the

$$\frac{\frac{\delta\rho c + \delta^2 c\left(1-\eta_1-\frac{\gamma}{2}\right)}{A^2} - \frac{\delta^2\gamma c}{2A^2}\frac{2\left[\delta\rho c + \delta^2 c\left(1-\eta_1-\frac{\gamma}{2}\right)\right]+\delta^2\gamma c}{2\left[\delta\rho c + \delta^2 c\left(1-\eta_2-\frac{\gamma}{2}\right)\right]+\delta^2\gamma c}}{\left[\frac{2\left[\delta\rho c + \delta^2 c\left(1-\eta_1-\frac{\gamma}{2}\right)\right]+\delta^2\gamma c}{2\left[\delta\rho c + \delta^2 c\left(1-\eta_2-\frac{\gamma}{2}\right)\right]+\delta^2\gamma c}\right]^{\frac{\gamma}{2-2\eta_2}}} \tag{I.1}$$

form in which the numerator is obviously a decreasing function of η_2. To yield the expected result in the easiest way, the denominator should be an increasing function of the same parameter. By directly differentiating the denominator with respect to η_2 and assuming that the differential is positive, we get that if the inequality

$$\ln\left[\frac{2\rho + 2\delta\left(1-\eta_1-\frac{\gamma}{2}\right)+\delta\gamma}{2\rho + 2\delta\left(1-\eta_2-\frac{\gamma}{2}\right)+\delta\gamma}\right] > -\frac{2\delta(1-\eta_2)}{2\rho + 2\delta\left(1-\eta_2-\frac{\gamma}{2}\right)+\delta\gamma} \tag{I.2}$$

holds, then $(h_1^*)_{\eta_2} > 0$. Now we can realize that the right-hand side expression in (I.2) is strictly increasing in η_2 hence

$$\sup_{\eta_2\in(0,0.5)}\left[-\frac{2\delta(1-\eta_2)}{2\rho + 2\delta\left(1-\eta_2-\frac{\gamma}{2}\right)+\delta\gamma}\right] = -\frac{\delta}{2\rho+\delta} < 0$$

which is over $\eta_2 = 0.5$. At the same time, for the left-hand side expression we have

$$\inf_{\eta_2\in(0,0.5)}\left\{\ln\left[\frac{2\rho + 2\delta\left(1-\eta_1-\frac{\gamma}{2}\right)+\delta\gamma}{2\rho + 2\delta\left(1-\eta_2-\frac{\gamma}{2}\right)+\delta\gamma}\right]\right\} = \ln\left[\frac{2\rho + 2\delta\left(1-\eta_1-\frac{\gamma}{2}\right)+\delta\gamma}{2\rho + 2\delta\left(1-\frac{\gamma}{2}\right)+\delta\gamma}\right]$$

which is over $\eta_2 = 0$ as the expression is strictly increasing in η_2. Now, the infimum for η_2 depends on η_1, in which the infimum is decreasing, so we can eventually say that

$$\inf_{\eta_1\in(0,0.5)}\left\{\inf_{\eta_2\in(0,0.5)}\left\{\ln\left[\frac{2\rho + 2\delta\left(1-\eta_1-\frac{\gamma}{2}\right)+\delta\gamma}{2\rho + 2\delta\left(1-\eta_2-\frac{\gamma}{2}\right)+\delta\gamma}\right]\right\}\right\} = \ln\left(\frac{\rho+\frac{\delta}{2}}{\rho+\delta}\right) < 0$$

which is over $\eta_1 = 0.5$. Therefore, as long as

$$g(\delta, \rho) \equiv \ln\left(\frac{\rho+\frac{\delta}{2}}{\rho+\delta}\right) + \frac{\delta}{2\rho+\delta} > 0, \tag{I.3}$$

$(h_1^*)_{\eta_2} > 0$ certainly holds. Over the point $(\delta, \rho) = (0,0)$, the function in (I.3) is not interpreted. In the δ-axis, where $\rho = 0$, $g(\delta, 0) = \ln(0.5) + 1 > 0$, while in the ρ-axis, where $\delta = 0$, $g(0, \rho) = \ln(1) + 0 = 0$.

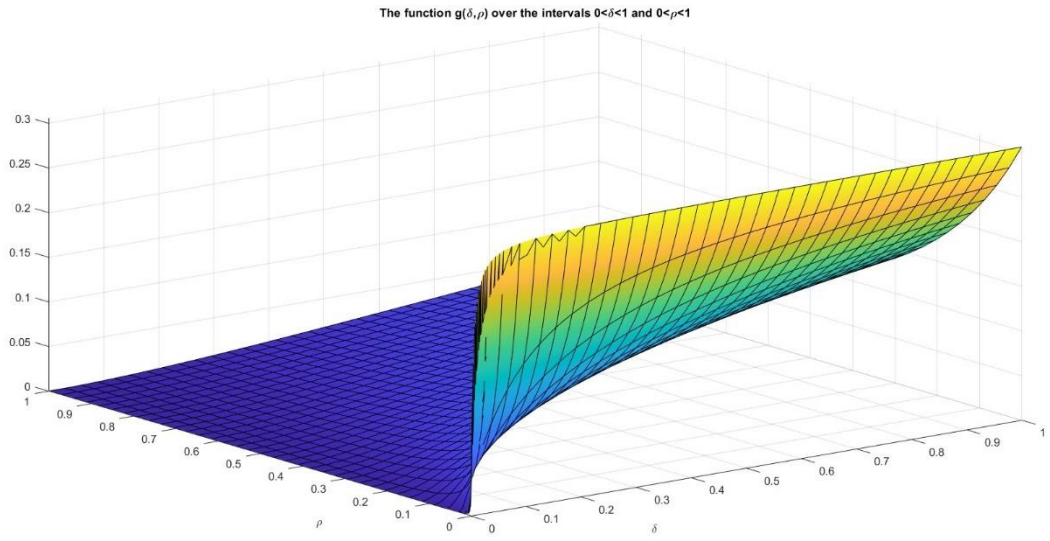


Figure 12: The function $g(\delta, \rho)$ over the intervals $0 < \delta < 1$ and $0 < \rho < 1$.

These cases, however, are irrelevant as some of the basic assumptions of the model are omitted (there is no discounting, or there is no capital depreciation). To see that $g(\delta, \rho)$ is strictly positive over the relevant domain of $\delta > 0$ and $\rho > 0$, suffice it to evaluate its differential

$$g'_\delta = \frac{\delta\rho}{(\delta+\rho)(\delta+2\rho)^2}$$

which is positive, so g is strictly increasing in δ as long as $\delta > 0$ at any $\rho > 0$. Over the relevant domain, $(h_1^*)_{\eta_2} > 0$ therefore holds.

□

Proposition 11: *The unique optimum capital stock* h_2^* *as determined by* $h(h_1^*)$ *in (29) and the term in (28) is an increasing function of* η_1*, therefore* $(h_2^*)_{\eta_1} > 0$.

Proof: We can follow the same strategy as in the case of proving $(h_1^*)_{\eta_2} > 0$. Starting from (25b), we get

$$(-1)(h_2^*)^{-2}(h_2^*)_{\eta_1} - A\left(1-\eta_2-\frac{\gamma}{2}\right)\left(\eta_2+\frac{\gamma}{2}\right)\theta_2^* u_2^* (h_1^*)^{\frac{\gamma}{2}}(h_2^*)^{\eta_2+\frac{\gamma}{2}-2}(h_2^*)_{\eta_1}$$
$$- A\left(1-\frac{\gamma}{2}\right)\frac{\gamma}{2}\theta_1^* u_1^* (h_1^*)^{\eta_1+\gamma}(h_2^*)^{\frac{\gamma}{2}-2}(h_2^*)_{\eta_1} + A\frac{\gamma}{2}\theta_1^* u_1^* (h_1^*)^{\eta_1+\frac{\gamma}{2}}(h_2^*)^{\frac{\gamma}{2}-1}\ln(h_1^*) = 0$$

which, like previously, is the sum of four terms, including three parameters on $(h_2^*)_{\eta_1}$. Under the stated assumptions $\gamma > 0$ and $0.5 > \eta_i + \frac{\gamma}{2} > 0$, and given the results that $h_i^* > 0$, $\theta_i^* > 0$ and $u_i^* > 0$, where $i = 1,2$, all the three parameters on $(h_2^*)_{\eta_1}$ are negative, while the fourth is positive, so after rearranging the equation we receive $(h_2^*)_{\eta_1} > 0$.

□

Proposition 12: *The unique optimum capital stock* h_1^* *as determined by* $h(h_1^*)$ *in (29) is an increasing function of* η_1, therefore $(h_1^*)_{\eta_1} > 0$.

Proof: By direct differentiation, the exponent on h_1^* in (29) turns out to be a decreasing function of η_1. Therefore, we only need to show that the multiplier in (29) is also a decreasing function of η_1. The differential of the multiplier with respect to η_1 is negative if the inequality

$$\frac{c\delta^3\gamma^2\left[\frac{\rho+\delta(1-\eta_1)}{\rho+\delta(1-\eta_2)}\right]^{\frac{\gamma}{2(\eta_2-1)}}}{4A^2(\eta_2-1)[\rho+\delta(1-\eta_1)]} < \frac{c\delta^2\left[\frac{\rho+\delta(1-\eta_1)}{\rho+\delta(1-\eta_2)}\right]^{\frac{\gamma}{2(\eta_2-1)}}(2\eta_2+\gamma-2)[2\rho+\delta(2-2\eta_2-\gamma)]}{4A^2(\eta_2-1)[\rho+\delta(1-\eta_2)]}$$

holds. We can proceed by assuming that the inequality holds to check if it leads to a contradiction. After manipulation we get the inequality

$$\frac{\delta\gamma^2}{\rho+\delta(1-\eta_1)} > \frac{(2\eta_2+\gamma-2)[2\rho+\delta(2-2\eta_2-\gamma)]}{\rho+\delta(1-\eta_2)}$$

in which the left-hand side expression is positive while the right-hand side expression is negative for $2\eta_2+\gamma-2<0$ under the parameter restrictions. Therefore, $(h_1^*)_{\eta_1} > 0$.

□

Proposition 13: *The unique optimum capital stock* h_2^* *as determined by* $h(h_1^*)$ *in (29) and the term in (28) is an increasing function of* η_2, *therefore* $(h_2^*)_{\eta_2} > 0$.

Proof: It has already been demonstrated that $(h_1^*)_{\eta_2} > 0$. The exponent on h_1^* in (28) is also an increasing function of η_2. As $h_1^* > 1$, the term

$$(h_1^*)^{\frac{1-\eta_1}{1-\eta_2}}$$

is an increasing function of η_2. Now, we only need to show that the multiplier in (28) is also an increasing function of η_2. The differential of the multiplier with respect to η_2 is positive if

$$\frac{\ln\left[\frac{\rho+\delta(1-\eta_1)}{\rho+\delta(1-\eta_2)}\right]+1}{2\left[\frac{\rho+\delta(1-\eta_1)}{\rho+\delta(1-\eta_2)}\right]^{\frac{1}{2(\eta_2-1)}}(\eta_2-1)^2} > \frac{\rho}{2[\rho+\delta(1-\eta_2)]\left[\frac{\rho+\delta(1-\eta_1)}{\rho+\delta(1-\eta_2)}\right]^{\frac{1}{2(\eta_2-1)}}(\eta_2-1)^2}$$

holds that implies the inequality

$$\ln\left[\frac{\rho+\delta(1-\eta_1)}{\rho+\delta(1-\eta_2)}\right]>\frac{\rho}{\rho+\delta(1-\eta_2)}-1$$

that certainly holds if the infimum of the left-hand side expression is greater than the supremum of the right-hand side expression. Playing on the monotonicity the outer function on the left, it is easy to realize that the infimum is there where the numerator has its infimum (which is at $\eta_1 = 0.5$) and the denominator has its supremum (which is at $\eta_2 = 0$). Consequently, the infimum of the left-hand side expression is

$$\ln\left[\frac{\rho+\frac{\delta}{2}}{\rho+\delta}\right].$$

To find the supremum of the right-hand side expression, we first realize that the expression is an increasing function of η_2, so its supremum is

$$-\frac{\frac{\delta}{2}}{\rho+\frac{\delta}{2}}$$

where $\eta_2 = 0.5$. Consequently, $(h_2^*)_{\eta_2} > 0$ since the inequality

$$\ln\left[\frac{\rho+\frac{\delta}{2}}{\rho+\delta}\right]+\frac{\delta}{2\rho+\delta}>0$$

certainly holds as we pointed out in the case of $(h_1^*)_{\eta_2} > 0$.

□

Appendix J

Proposition 14: *The Hessian $\widetilde{H}$ of $\tilde{J}^*(\beta)$ is indefinite.*

Proof: We must take into account the fact that $\widetilde{H}$ containing the second partial differentials of $\tilde{J}^*(\beta)$ with respect to the parameters A, η_1 and η_2 is certainly symmetric. Therefore, it can be rewritten in the

$$\widetilde{H}=\begin{bmatrix}0 & a & b\\ a & c & 0\\ b & 0 & d\end{bmatrix}$$

form with all nonzero elements being strictly positive and real. From this $\left|\widetilde{H}\right| = -a^2d - b^2c < 0$. Given that the determinant is the product of the eigenvalues, it follows that $\widetilde{H}$ is indefinite (in the case where one eigenvalue is negative while the other two are positive) or negative definite (in the case where all three eigenvalues are negative). For a matrix to be negative definite, Sylvester's alternating sign condition must hold (Horn & Johnson, 2013, pp. 430-439) according to which

$$(-1)^1 D_1 > 0$$
$$(-1)^2 D_2 > 0$$
$$(-1)^3 D_3 > 0$$

where $D_1 < 0$, $D_2 > 0$ and $D_3 < 0$ are the leading principal minors alternating signs. In the case of $\widetilde{H}$, $D_1 < 0$ certainly does not hold as $D_1 = 0$. Likewise, $D_2 > 0$ does not hold either as $D_2 = -a^2 < 0$. The criterion is conclusive here even though the first leading principal minor that breaks the pattern of Sylvester's criterion for negative definiteness is zero (i.e. $\widetilde{H}_{11} = D_1 = 0$). On this ground, $\widetilde{H}$ might also be negative semidefinite. Negative semi-definiteness, however, is ruled out by the strict negativity of $|\widetilde{H}|$. Under negative semi-definiteness, at least one eigenvalue would be zero and the determinant would also be zero. Therefore, $\widetilde{H}$ is indefinite.

□

References

Barro, R. J. (1974). Are government bonds net wealth? *Journal of Political Economy, 82*(6), 1095-1117.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy, 75*(4), 352-365.

Bertsekas, D. P. (1976). *Dynamic programming and stochastic control.* New York: Academic Press.

Bilstein, R. E. (1980). *Stages to Saturn. A technological history of the Apollo/Saturn Launch Vehicles.* Washington, DC: NASA History Office.

Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement. A meta-analysis. *Review of Educational Research, 73*(2), 125-230.

Caputo, M. R. (1990a). A qualitative characterization of the competitive nonrenewable resource extracting firm. *Journal of Environmental Economics and Management, 18*(3), 206-226.

Caputo, M. R. (1990b). How to do comparative dynamics on the back of an envelope in optimal control theory. *Journal of Economic Dynamics and Control, 14*(3-4), 655-683.

Caputo, M. R. (2005). *Foundations of dynamic economic analysis. Optimal control theory and applications.* Cambridge: Cambridge University Press.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity.* Washington, D.C.: Office of Education, U.S. Department of Health, Education & Welfare.

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica, 78*(3), 883-931.

Eisner, R., & Strotz, R. H. (1962). The theoretical framework. In R. Eisner, & R. H. Strotz, *Determinants of business investment* (pp. 61-116). Englewood Cliffs, NJ: Prentice Hall.

Epple, D., & Romano, R. E. (1998). Competition between private and public schools, vouchers, and peer-group effects. *The American Economic Review, 88*(1), 33-62.

Gutierrez, S., Whitmer, R. A., Soh, Y., Peterson, R., George, K. M., Lor, Y., . . . Gilsanz, P. (2024). School-based racial segregation, social support, and late-life cognitive function in the Study of Healthy Aging in African Americans (STAR). *Alzheimer's & Dementia, 20*(9), 6257-6267.

Heckman, J. J. (2009). The case for investing in disadvantaged young children. In B. Lesley (Ed.), *Big ideas for children. Investing in our nation's future* (pp. 49-58). Washington, DC: First Focus.

Heckman, J. J., & Masterov, D. V. (2007). The productivity argument for investing in young children. *Review of Agricultural Economics, 29*(3), 446–493.

Horn, R. A., & Johnson, C. A. (2013). *Matrix analysis* (2nd ed.). Cambridge, UK: Cambridge University Press.

Hoxby, C. M., & Weingarth, G. (2005). *Taking race out of the equation. School reassignment and the structure of peer effects. Mimeograph.* Retrieved from https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=386de4e83235738ec333387e23abf71a4c1f1e4d

Imberman, S., Kugler, A. D., & Sacerdote, B. (2009). *Katrina's children. Evidence on the structure of peer effects from hurricane evacuees. NBER Working Paper #15291.* Cambridge, MA: National Bureau of Economic Research.

Lavy, V., Silva, O., & Weinhardt, F. (2012). The good, the bad, and the average. Evidence on ability peer effects in schools. *Journal of Labor Economics, 30*(2), 367-414.

Lucas, R. E. (1965a, February). Optimal investment policy and the flexible accelerator. *International Economic Review, 8*(1), 78-85.

Lucas, R. E. (1965b). Distributed lags and optimal investment policy. In R. E. Lucas, & T. J. Sargent (Eds.), *Rational expectations and econometric practice* (pp. 39-54). Minneapolis: University of Minnesota Press.

Lucas, R. E. (1965c). Tests of a capital-theoretic model of technological change. *The Review of Economic Studies, 34*(2), 175-189.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics, 22*(1), 3-42.

Lucas, R. E., & Moll, B. (2014, February). Knowledge growth and the allocation of time. *Journal of Political Economy, 122*(1), 1-51.

Manski, C. F. (1993). Identification of endogenous social effects. The reflection problem. *The Review of Economic Studies, 60*(3), 531-542.

Neal, D. A. (2018). *Information, incentives, and education policy.* Cambridge, MA: Harvard University Press.

Oniki, H. (1973). Comparative dynamics (sensitivity analysis) in optimal control theory. *Journal of Economic Theory, 6*(3), 265-283.

Orfield, G., & Lee, C. (2007). *Historic reversals, accelerating resegregation, and the need for new integration strategies.* Los Angeles: Civil Rights Project/Proyecto Derechos Civiles, UCLA.

Peterson, R. L., George, K. M., Barnes, L. L., Gilsanz, P., Mayeda, E. R., Glymour, M. M., . . . Whitmer, R. A. (2021). Association of timing of school desegregation in the United States with late-life cognition in the Study of Healthy Aging in African Americans (STAR) Cohort. *JAMA Network Open, 4*(10), 1-10.

Romer, P. M. (1986). Increasing returns and long-run growth. *The Journal of Political Economy, 94*(5), 1002-1037.

Sacerdote, B. (2011). Peer effects in education. How might they work, how big are they and how much do we know thus far? In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. 3, pp. 249-277). Amsterdam: North Holland.

Weathers, E. S., & Sosina, V. E. (2022). Separate remains unequal. Contemporary segregation and racial disparities in school district revenue. *American Educational Research Journal, 59*(5), 905-938.