MPRA
Munich Personal RePEc Archive

Children Versus Ideas: an "Influential" Theory of Demographic Transition

Zakharenko, Roman

State University - Higher School of Economics (Moscow)

23 September 2009

Online at https://mpra.ub.uni-muenchen.de/17478/
MPRA Paper No. 17478, posted 23 Sep 2009 16:50 UTC

Children Versus Ideas: an “Influential” Theory of Demographic Transition

Roman Zakharenko*

September 23, 2009

Abstract

In this paper, I build on (Blackmore 2000) to propose a formal theory of demographic transition (fertility decline) and associated growth of the stock of knowledge. The novelty of this theory is to entirely exclude private consumption from the objective function of the decision makers, and to assume that their goal is to maximize their social *influence*, that is, the number of people in the next generation utilizing their ideas. With high communication costs, one's ideas are utilized mainly by his/her children, which creates an incentive to have as many children as possible. With modern communication technologies, one's ideas can be used by millions, which makes people invest time into improvement of own ideas rather than production of children. Even those who can influence only their own children are induced to have smaller families and improve own ideas, because their children now have access not only to ideas of parents but also to ideas from the outside world.

Keywords: demographic transition, social influence, economic growth
JEL codes: J13, O15

*State University - Higher School of Economics, Moscow, Russia. Email: r.zakharenko@gmail.com. Web: www.rzak.ru. I thank Maxim Arnold, Ken Binmore, and Lesley Newson, who have influenced this research.

1 Motivation

In the last ten or fifteen years, demographers have increasingly argued that one important determinant of fertility decline is *social influence* exerted on high-fertility societies by the outside world. People who have been exposed to the ideas delivered by the outside world are choosing to have smaller families. For example, (Axinn and Barber 2001) discover that proximity of young girls in rural Nepal to a school, controlling for their subsequent attendance of the school and their future husbands' attendance of the school, has a negative impact on their subsequent fertility. (Axinn and Yabiku 2001), from the same data, discover that proximity to such public places as a market or a bus stop during childhood increase the rate of contraception among women.

A good example of the opposite phenomenon – little contact with the outside world and high fertility – is Amish community in the United States. They voluntarily abstain from modern communication devices such as phones and television, as well as from modern vehicles and travel opportunities. Such an abstention effectively limits the number of people with whom Amish community members can interact, and increase the frequency of interaction with parents and other close relatives. At the same time, Amish families are very large. (Greksa 2002) estimates fertility among Old Order Amish at 7.7 children per woman, which makes Amish community one of the fastest-growing communities in the world, and certainly the fastest-growing community among the OECD countries.

In this paper, I propose a formal theory of demographic transition (fertility decline) that relates the outside social influence with fertility choices of the decision makers. Conventional economic theories of demographic transition are ill-equipped to modeling *social influence*: it is usually assumed that economic agents rationally maximize utility based on personal consumption and, occasionally, include children's consumption with a smaller weight (e.g. (Becker 1960), (Becker, Murphy and Tamura 1990), (Caldwell 1976), (Galor and Moav 2002)). The novelty of my theory is to abstract from consumption-based utility, and to embed social influence directly into individuals' utility function. The next paragraph outlines the theory in greater detail.

Every agent creates two types of output over his lifetime: *ideas* and children, each of which require a time input to produce. Since time is a limited resource, there exists a tradeoff between the quality of ideas and the number of children produced by an agent. Production of ideas requires absorption of ideas created by the previous generation; the goal of agents is to maximize their social influence, that is, the extent to which their own ideas have been absorbed by the next generation. Higher-quality ideas are absorbed more intensively, which creates an incentive to spend more time on development of ideas and have fewer children. Having more children, however, increases the number of people in the next generation utilizing one's ideas, which creates incentives to have more children and spend less time on development of ideas.

In traditional societies, non-relatives are isolated from each other, and children absorb ideas of their own parents regardless of the quality of these ideas. This creates incentives for parents to save time on development of ideas and have as many children as possible. In modern societies, learning from non-parents (and therefore teaching non-children) is nearly as easy as learning from parents, which makes demand for one's ideas more elastic and induces people to spend more time on development of their own ideas and thus have fewer children. In an extension of the model, I show that fertility declines even when people cannot influence the outside world themselves, but their children are influenced by the outside world. Too fertile, and therefore too uneducated, parents will not be able to influence even their own children in the presence of ideological competitors from the outside world.

1.1 Social influence versus social learning

While the effect of informational exposure to the outside world on fertility causes little doubt among demographers, the precise mechanism of how the former affects the latter is less obvious. This paper models the social influence mechanism, that is, families exposed to the outside world are less willing to have children. An alternative mechanism is *social learning*, according to which high fertility in traditional societies occurs not because parents are willing to have many children, but because they do not know how to control fertility. Interaction with the outside world brings knowledge of contraception techniques, and thus lowers fertility.

To find out which of the two factors – social influence or social learning – is significant in practice, (Behrman, Kohler and Watkins 2002) conduct a detailed empirical analysis of rural households in Kenya, and find that both play a significant role; social learning is relatively more important in more developed (urbanized) areas, while social influence plays a role in more remote places. This finding can be interpreted as follows: in urban areas, the *willingness* to have smaller families emerged before *ability* to control fertility; thus, fertility declines due to acquisition of new "skills" rather than changes in preferences over the family size. In more remote places, it is not (in)ability but (un)willingness to control family size that is most important for fertility.

An additional evidence for this argument comes from (Bongaarts and Watkins 1996): they show that countries that entered demographic transition fairly late for their level of development (e.g. Latin American countries), had experienced faster decline of fertility than other countries. By the time of the beginning of the fertility decline, Latin American countries apparently had already been influenced by the Western world and had unfulfilled demand for small families; social learning allowed them to rapidly meet their demand. In less developed countries, fertility declines more slowly, apparently because it is driven not by social learning but by social influence which operates more slowly.

Thus, we may conclude, while social learning does play a role in fertility decline, social influence is a more fundamental factor of fertility change. After all, Amish community members who live in the middle of the most advanced civilization have easy access to all contraception techniques they need; their large family sizes are determined by (lack of) social influence, rather than lack of knowledge about contraception.

1.2 The willingness to influence: evidence and theoretical background

One of key innovations of this paper is to assume that people care about their social influence rather than their consumption of goods and services. In this section, I justify this assumption.

There is plenty of anecdotal evidence that many people care about their own lifetime legacy being utilized by the next generations. In pre-industrial societies, many fathers wanted their sons to continue their family business. Many elderly people want their values to be absorbed by younger generations. There are countless examples of people giving up their wealth, or even risking their lives, for the sake of defending and popularizing their ideology. Most scientists want their research to be cited, even when their own wealth no longer depends on the number of citations. Suicide bombers sacrifice their lives in exchange for high respect in their community.

Modern Economics originates from Adam Smith who defined the objective of people as the maximization of personal wealth; Adam Smith himself, however, did not seem to maximize his wealth: writing "The Wealth of Nations " took twelve years of Smith's life, between age 41 and 53, and was probably in contemplation twelve years prior to that (Rae 2006 (1895)). Thus, developing his theory took a quarter of Smith's life; the monetary reward for his effort was 300 pounds (Rae 2006 (1895)), which was equivalent to his yearly income. Moreover, given eighteenth-century life expectancy, it was not obvious at all that Smith will be able to complete his project and receive his monetary reward. Thus, we may conclude, engaging is such an extensive project was probably not the first-best income maximization strategy. Adam Smith's behavior is rational if his objective was to maximize his social influence (in which he was highly successful); if he was a selfish *Homo economicus*, he should probably not have written his book.

Why is the desire to influence so strong that it may cause people to sacrifice their biological fitness? Below, I outline an evolutionary explanation. For most of their evolutionary history, humans lived in small groups of less than one hundred people, with very limited contact between the groups (Cosmides and Tooby 1994). Evolutionary psychology suggests that human preferences must have been "hard-wired" to living in such environment (Tooby and Cosmides 1992). In such traditional societies, highly regarded individuals were likely to become patriarchs and matriarchs (Newson, Postmes, Lea and Webley 2005), that is, their ideological fitness could have been positively correlated with biological fitness, which could cause an emergence of preference for social influence. Modern environment, however, is very different from the environment in which human psychology has evolved: thanks to travel possibilities and mass-media, humans routinely interact, personally and impersonally, with thousands of people; some celebrities influence the minds of millions. In this new environment, the preference for social influence is maladaptation from the standpoint of genetic fitness: the desire to become a patriarch with many children has been replaced by the desire to become a top manager, or a movie star, or a top scientist, or a politician, which requires having

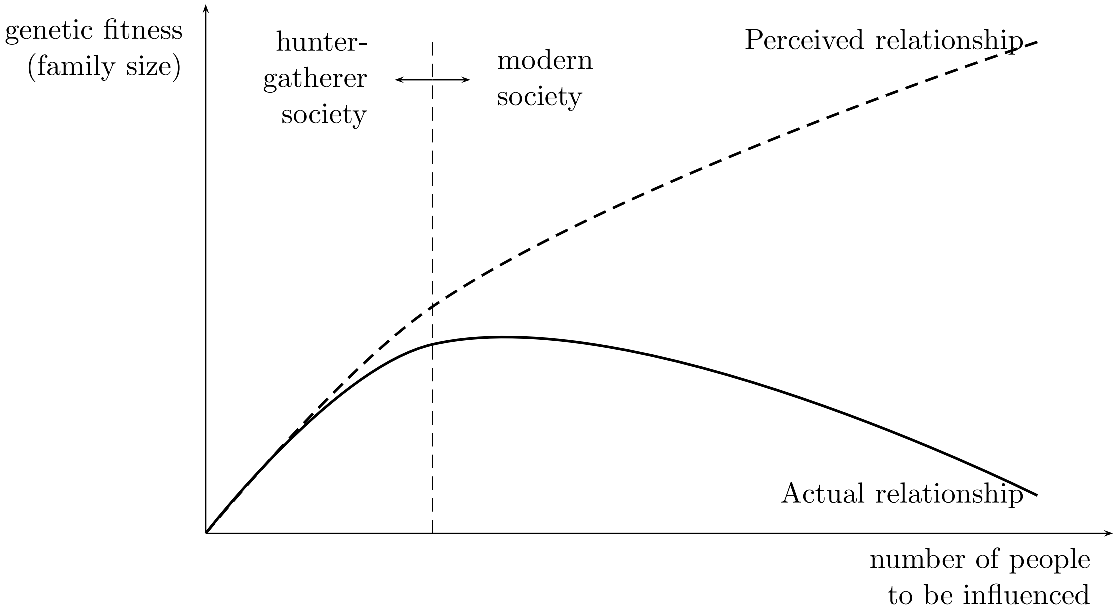

Figure 1: Preference for social influence: an illustration

a small family and spending a lot of time at work. Figure 1 illustrates the argument.

Two influential theories – memetics (Dawkins 2006 (1976)), (Blackmore 2000) and dual inheritance theory (Boyd and Richerson 1985) make a clear distinction between humans and all other species: humans are unique in their ability to imitate one another's behavior and communicate information. Only humans can directly copy goals and motor patterns of each other (Henrich and Gil-White 2001); in the language of this paper, only humans are able to absorb each other's ideas. Thus, humans are the only species that are able to produce not only biological, but also ideological offspring; therefore, humans are the only species that may theoretically deviate from standard Darwinian incentive of producing as many biological offspring as possible.

Dual inheritance theory does not suggest that being imitated is the objective of humans; imitation of the most successful peers is a skill that benefits the imitators, but has no effect on those being imitated. Memetics assumes that humans brains are infected by "viruses of the mind", or *memes*, whose only goal is to get copied into other brains; the role of humans is diminished to simply hosting the memes and doing what is best for them. This research is similar to memetics in the sense that humans are willing to seed their ideas into as many brains as possible; however, I abstract from empirically unobserved memes and assume that humans themselves desire to spread their ideas. Otherwise, the ideology of this research is similar to that of (Blackmore 2000) who offers a "memetic" theory of demographic transition: in isolated communities, memes cannot go far, and the "standard" genetic incentive to have as many children as possible dominates; small-family communities simply become extinct. When long-distance communication becomes available, people spend more time sending out their memes and thus have smaller families.

2 Model

2.1 Basics

The environment is characterized by non-overlapping generations[1] indexed by t. Each generation t is populated by a continuum G_t of agents.

Members of two consecutive generations are linked to each other by parent-child relationship: each child has one parent.[2] Formally, I define a parental operator $\mathcal{C} : G_{t-1} \to G_t$ such that for every $i \in G_{t-1}$, $\mathcal{C}(i)$ is the set of all children of i.

2.2 Constraints

I describe an individual's problem in a somewhat unusual way: I first present their constraints and then define the utility function. Each individual $j \in G_t, \forall t$ is endowed with $\overline{L}$ units of time that can be separated into two activities: learning L_j and production of children. I assume that raising each child takes ν units of parent's time, so an individual j who has n_j children has the following budget constraint:

$$L_j + \nu n_j \leq \overline{L} \tag{1}$$

Note that individuals are allowed to choose non-integer number of children, which can be interpreted as follows. While the choice of parents n_j may be non-integer, the actual number of children is a multinomial random variable with an expectation of n_j. As defined below, parents are risk neutral with respect to the number of children, and thus the variance in the number of children is immaterial: any mean-preserving spread in the distribution of the number of children does not change parents' decisions. Also, I assume that the time burden of raising children depends on expected value n_j, rather than on actual realization.

Each individual $j \in G_t$ is characterized by the quality q_j of his *idea*. In this stylized model, we assume for simplicity there is only one idea per person, which represents all intellectual legacy produced by that individual over his or her lifetime. To develop an idea, j has to absorb ideas developed by the previous generation. This assumption relies on the fact that knowledge accumulation is a social activity; the most successful ideas "stand on the shoulders of giants" rather than are developed in isolation. The parent of j has a strictly positive influence on the formation of j's idea, while everybody else in generation $t-1$ (denote them *non-parents*) have an infinitesimal impact on j. Given that there is a continuum of non-parents, all of them together have a sizeable

[1]It is more intuitive to assume that generations do overlap, so that younger generation could absorb ideas of living members of the older generation. Non-overlapping generations structure was introduced to simplify exposition of the model.

[2]I abstract from genders in the model; see (Becker et al. 1990) and (Galor and Moav 2002) for similar assumptions

impact on the formation of j's idea. Formally,

$$q_j = \left(M \left(q_{\mathcal{C}^{-1}(j)} x_{\mathcal{C}^{-1}(j)j} \right)^\gamma + \int_{i \in G_{t-1} \setminus \mathcal{C}^{-1}(j)} \left(q_i x_{ij} \right)^\gamma di \right)^{\frac{1}{\gamma}} \tag{2}$$

Here M is a positive constant which shows the weight of a parent, relative to non-parents, in the formation of one's idea; $\mathcal{C}^{-1}(j)$ indicates the parent of j; $x_{ij} > 0$ is the *intensity* of j's learning from $i \in G_{t-1}$: it indicates to what extent individual j has absorbed i's ideas. While q_i are given for the person j, he is free to choose his learning intensities x_{ij} subject to a time constraint outlined below.

Absorbtion of ideas incurs a time cost of τ_{ij} per unit of learning intensity. The total amount of time spent on learning from all sources must add up to L_j:

$$M x_{\mathcal{C}^{-1}(j)j} \tau_{\mathcal{C}^{-1}(j)j} + \int_{i \in G_{t-1} \setminus \mathcal{C}^{-1}(j)} x_{ij} \tau_{ij} di = L_j \tag{3}$$

I assume that the time cost of learning from one's parent is fixed and is normalized to unity:

$$\tau_{\mathcal{C}^{-1}(j)j} \equiv 1, \forall j \in G_t, \forall t \tag{4}$$

The time cost of learning from non-parents may vary; the main goal of this research is to study the effect of changes in the cost of learning from non-parents on family size.

2.3 Objective function

Finally, I define the objective function of individuals. I assume that individuals are willing to maximize the extent to which their idea has been utilized by the next generation; they place a special emphasis on their own children. More formally, the extent to which j's child has absorbed his/her ideas has a strictly positive weight on j's utility, while each of non-children has an infinitesimal impact. Since there is a continuum of non-children, all of them combined have a sizeable effect on j's utility:

$$U_{j \in G_t} = n_j M x_{j\mathcal{C}(j)} + \int_{k \in G_{t+1} \setminus \mathcal{C}(j)} x_{jk} dk \tag{5}$$

Given that all children of any given j are identical, by $x_{j\mathcal{C}(j)}$ I mean, with a slight abuse of notation, the intensity of learning from j by his representative child. Note that n_j is (generically) non-integer *expected* number of children, hence the above formula describes the expected utility rather than its actual realization.

Note that individual j does not choose the learning intensities of the next generation x_{jk}; the control parameters of j are the learning time L_j, the number of children n_j, and his own learning intensities x_{ij} for $i \in G_{t-1}$.

3 Analysis

3.1 Optimal learning intensities

It is intuitively obvious and is formally shown below that the intensity of learning x_{jk} of $k \in G_{t+1}$ from $j \in G_t$ is increasing with the quality q_j of j's idea; therefore, for any given value of learning time L_j individual j attempts to achieve the highest possible q_j. This allows us to find the optimal learning intensities x_{ij} from every teacher i. By maximizing (2) subject to (3) over all $x_{ij}, \forall i \in G_t$, we come up with the optimal learning intensities (see appendix A for derivation). Before presenting the solution, it is convenient to define the following variable:

$$E_j \equiv M \left(q_{\mathcal{C}^{-1}(j)}\right)^{\frac{\gamma}{1-\gamma}} + \int_{i \in G_{t-1} \setminus \mathcal{C}^{-1}(j)} \left(\frac{q_i}{\tau_{ij}}\right)^{\frac{\gamma}{1-\gamma}} di \tag{6}$$

The variable E_j describes the *learning environment* of individual j: it shows how smart j's potential teachers are, and how easy it is to absorb their ideas. By definition, E_j does not depend on j's decisions and is taken by him as given.

Now, we are equipped to demonstrate the optimal learning intensities:

$$x_{ij} = \frac{L_j}{E_j} q_i^{\frac{\gamma}{1-\gamma}} \tau_{ij}^{-\frac{1}{1-\gamma}} \tag{7}$$

Learning intensity of j from i increases linearly with the overall length of education L_j; it increases with the quality of teacher's idea q_i (as claimed above), decreases with learning cost τ_{ij}. Learning intensity from any given teacher also decreases as the overall learning environment E_j improves.

From (2) and (7), we can also compute the optimal (highest possible) quality of j's idea given the learning time L_j:

$$\begin{aligned} q_j &= \frac{L_j}{E_j} \left(M \left(q_{\mathcal{C}^{-1}(j)}\right)^{\frac{\gamma}{1-\gamma}} + \int_{i \in G_{t-1} \setminus \mathcal{C}^{-1}(j)} \left(\frac{q_i}{\tau_{ij}}\right)^{\frac{\gamma}{1-\gamma}} di \right)^{\frac{1}{\gamma}} = \frac{L_j}{E_j} E_j^{\frac{1}{\gamma}} \\ &= E_j^{\frac{1-\gamma}{\gamma}} L_j \end{aligned} \tag{8}$$

Using (8), we can rewrite (7) and (6):

$$x_{ij} = \frac{L_j}{E_j} E_i L_i^{\frac{\gamma}{1-\gamma}} \tau_{ij}^{-\frac{1}{1-\gamma}} \tag{9}$$

$$E_j = M E_{\mathcal{C}^{-1}(j)} \left(L_{\mathcal{C}^{-1}(j)}\right)^{\frac{\gamma}{1-\gamma}} + \int_{i \in G_{t-1} \setminus \mathcal{C}^{-1}(j)} E_i \left(\frac{L_i}{\tau_{ij}}\right)^{\frac{\gamma}{1-\gamma}} di \tag{10}$$

Using (1) and (9), we can rewrite (5) as follows:

$$U_j = L_j^{\frac{\gamma}{1-\gamma}} E_j \left[M \frac{\overline{L} - L_j}{\nu} \frac{L_{\mathcal{C}(j)}}{E_{\mathcal{C}(j)}} + \int_{k \in G_{t+1} \setminus \mathcal{C}(j)} \frac{L_k}{E_k} \tau_{jk}^{-\frac{1}{1-\gamma}} dk \right] \tag{11}$$

The objective of individual $j \in G_t$ is to maximize (11) over L_j, keeping in mind that $E_k, k \in G_{t+1}$ depend on L_j. The relationship between the latter two can be described by extrapolating (10) one generation forward:

$$E_k = M E_{\mathcal{C}^{-1}(k)} L_{\mathcal{C}^{-1}(k)}^{\frac{\gamma}{1-\gamma}} + \int_{j' \in G_t \setminus \mathcal{C}^{-1}(k)} E_{j'} \left(\frac{L_{j'}}{\tau_{j'k}}\right)^{\frac{\gamma}{1-\gamma}} dj' \tag{12}$$

When k is a child of j ($k \in \mathcal{C}(j)$), the above formula simplifies to

$$E_{\mathcal{C}(j)} = M E_j L_j^{\frac{\gamma}{1-\gamma}} + \int_{j' \in G_t \setminus j} E_{j'} \left(\frac{L_{j'}}{\tau_{j'k}}\right)^{\frac{\gamma}{1-\gamma}} dj' \tag{13}$$

When k is a child of j, L_j has a sizeable effect on k's learning environment; otherwise, the effect of L_j on E_k is negligible: j is one of continuum of non-parents of k that affect it.

3.2 Symmetric learning environment

Definition 1 *Symmetric learning environment occurs when all members of any given generation have the same learning environment:* $E_i = E_j \equiv E_t, \forall i \in G_t, \forall j \in G_t, \forall t$

In the rest of the paper, we will only deal with symmetric learning environment. For this environment, the following several results can be established.

Proposition 1 *In symmetric learning environment, one's utility does not depend on his or her learning environment*

Intuitively, a better learning environment (smarter ancestors or easier access to the knowledge of ancestors) makes not only $j \in G_t$, but also all his contemporaries smarter; as a result, the influence of j on the next generation remains the same as before.

Proof. Consider the learning environment of an agents $k \in G_{t+1}$ shown in (12). It depends on learning environments of agents from the previous generation t. If the latter have the same learning environments, they can be factored out in (12):

$$E_{t+1} \equiv E_k = E_t \left(M L_{\mathcal{C}^{-1}(k)}^{\frac{\gamma}{1-\gamma}} + \int_{j' \in T(k)} \left(\frac{L_{j'}}{\tau} \right)^{\frac{\gamma}{1-\gamma}} dj' \right), \forall k \in G_{t+1} \tag{14}$$

Now, consider the utility of an agent $j \in G_t$ shown in (11). It is proportional to his own learning environment $E_j \equiv E_t$. It also depends on learning environments of the next generation $E_k, k \in G_{t+1}$, which, by assumption, are equal to each other and thus can be factored out in the expression for utility:

$$U_j = L_j^{\frac{\gamma}{1-\gamma}} \frac{E_t}{E_{t+1}} \left[M \frac{\overline{L} - L_j}{\nu} L_{\mathcal{C}(j)} + \int_{k \in G_{t+1} \setminus \mathcal{C}(j)} L_k \tau_{jk}^{-\frac{1}{1-\gamma}} dk \right] \tag{15}$$

where E_{t+1} is the learning environment of some representative $k \in G_{t+1}$. From (14), E_{t+1} is proportional to E_t, which means that the latter cancels out from the expression for utility (15). Thus, one's utility does not depend on his learning environment. ■

Independence of utility on one's learning environment allows us to establish the following important result.

Proposition 2 *Members of generation t, when making optimal learning choices, treat the learning choices of the next generation as given. Formally,* $\frac{dL_k}{dL_j} = 0, \forall j \in G_t, \forall k \in G_{t+1}, \forall t$

Proof. From Proposition 1, it follows that the utility of $j \in G_t$, and thus his optimal learning time L_j, do not depend on his learning environment: $\frac{dL_j}{dE_j} = 0$. By extrapolating this result one generation forward, we get $\frac{dL_k}{dE_k} = 0, \forall k \in G_{t+1}$. Therefore, learning decisions of j, which affect the learning environment of k, do not affect the learning decisions of the latter. Formally,

$$\frac{dL_k}{dL_j} = \frac{dL_k}{dE_k} \frac{dE_k}{dL_j} = 0 \tag{16}$$

While $\frac{dE_k}{dL_j}$ is positive, the effect of L_j on L_k is zero. ■

Without this result, we would need to assume that L_k is an unknown function of L_j, and solve a differential equation to find the optimal value of the latter. When L_k does not depend on L_j, solving an agent's problem becomes a trivial task, with closed-form solutions in many cases.

4 Results

4.1 A primer: isolated families

Consider an extreme case in which children can only learn from their parents, that is $\tau_{ij} = \infty, \forall i \neq \mathcal{C}^{-1}(j)$. Then, utility (11) simplifies to

$$U_j = L_j^{\frac{\gamma}{1-\gamma}} E_j M \frac{\overline{L} - L_j}{\nu} \frac{L_{\mathcal{C}(j)}}{E_{\mathcal{C}(j)}} \tag{17}$$

where $E_{\mathcal{C}(j)}$ takes a simple form of

$$E_{\mathcal{C}(j)} = M E_j L_j^{\frac{\gamma}{1-\gamma}} \tag{18}$$

By plugging (18) into (17), we end up with the following optimization problem:

$$\max_{L_j} \frac{\overline{L} - L_j}{\nu} L_{\mathcal{C}(j)} \tag{19}$$

From Proposition 2, $L_{\mathcal{C}(j)}$ does not depend on j's decisions. Then, it is straightforward to observe that individual j chooses the lowest possible level of education $L_j = \underline{L}$. Intuitively, when children can learn only from their parents, the latter are effectively monopolists in the ideological market for their children and have no incentives to improve the quality of their own ideas. Parents choose to maximize the number of their children: having secured the *share* of their ideological market at 100%, they maximize the *size* of that market.

Therefore, when families have limited ability to communicate with one another, we end up with a standard Darwinian/Malthusian equilibrium in which parents maximize the number of children and have the minimal level of education.

4.2 Learning from parent's neighbors

For further analysis, it is convenient to define a spatial allocation of agents, which enables us to measure distance between any two agents and define a "neighborhood" of one's parent. I assume that all members of a given generation are distributed uniformly on a circle. The density of agents on the circle is constant, so the circle expands with population growth.

To define the structure of the learning costs, I divide the set G_{t-1} of all potential teachers of any given individual $j \in G_t$ of any given generation t into three subsets. The first subset consists of only one element, the parent of j; the cost of learning from the parent is unity as before: $\tau_{\mathcal{C}^{-1}(j)j} = 1$. The second subset is the mass N_0 of the "neighbors" of $\mathcal{C}^{-1}(j)$, denote it $T(j)$ (where T stands for "teachers"). Formally, $T(j)$ is a subset of $G_{t-1} \backslash \mathcal{C}^{-1}(j)$ such that $|T(j)| = N_0$ and $||i, \mathcal{C}^{-1}(j)|| \leq ||i', \mathcal{C}^{-1}(j)||$ for any $i \in T(j)$ and $i' \in G_{t-1} \backslash T(j) \backslash \mathcal{C}^{-1}(j)$. The operator $|X|$ here is the mass ("number") of all agents in the set X, while $||x, y||$ measures distance between elements x

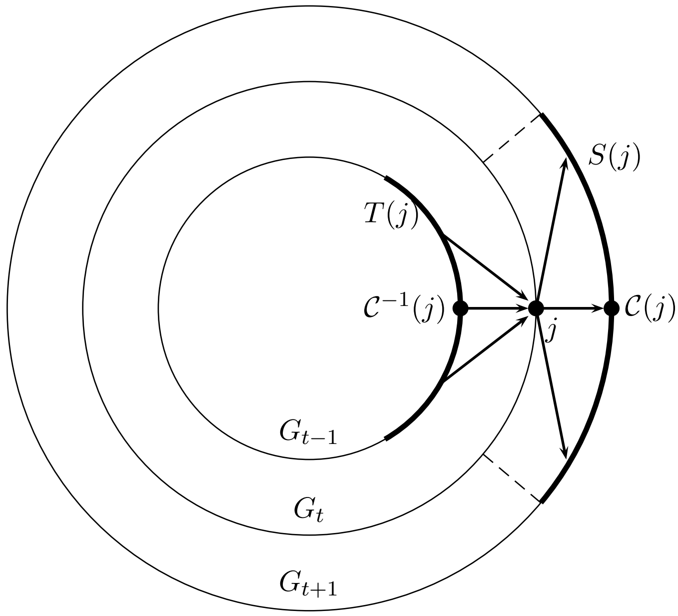

Figure 2: The geography of influence: an illustration

and y of a set. I assume that the cost of learning from parent's neighbors is $\tau_{i,j} = \tau \geq 1, \forall i \in T(j)$

Finally, the third subset of all potential teachers is "non-neighbors" of j's parent; I assume that learning from them is impossible. Formally,

$$\tau_{i,j} = \infty, \forall i \in G_{t-1} \backslash T(j) \backslash \mathcal{C}^{-1}(j)$$

For any given $j \in G_t$, define $S(j)$ as a set of all potential students of j, besides j's own children, who's learning cost from j is finite. Formally, $S(j)$ is the set of all $k \in G_{t+1}$ such that $j \in T(k)$. Assuming that the representative family size (number of children) is n_t, the measure of the set $S(j)$ is $n_t N_0$ for all $j \in G_t$, for all t.

The key variable of interest here is N_0, the mass of neighbors of one's parent from which it is possible to learn: it shows the extent to which people are exposed to the outside world. Traditional societies, in which people interact with a small number of neighbors who live in the same village, correspond to a small value of N_0. Modern world, where most people live in large cities and frequently travel to other cities, corresponds to large N_0. Below, we analyze how family size decisions depend on N_0. An extreme case of $N_0 = 0$, when learning from parent's neighbors is impossible and families are virtually isolated from each other, has been analyzed in section 4.1.

Due to symmetry, all agents $k \in G_{t+1}$ choose the same learning intensity $L_k = L_{t+1}, \forall k$ which can therefore be factored out in the expression for utility (11). Then, j's utility can now be rewritten as

$$U_j = L_{t+1} L_j^{\frac{\gamma}{1-\gamma}} E_t \left[M \frac{\overline{L} - L_j}{\nu} \frac{1}{E_{\mathcal{C}(j)}} + \int_{k \in S(j)} \frac{1}{E_k} \tau^{-\frac{1}{1-\gamma}} dk \right]$$
$$= L_{t+1} L_j^{\frac{\gamma}{1-\gamma}} E_t \left[M \frac{\overline{L} - L_j}{\nu} \frac{1}{E_{\mathcal{C}(j)}} + n_t N_0 \frac{1}{E_{t+1}} \tau^{-\frac{1}{1-\gamma}} \right] \tag{20}$$

where $E_{\mathcal{C}(j)}$ can be rewritten from (13) as

$$E_{\mathcal{C}(j)} = E_t \left(M L_j^{\frac{\gamma}{1-\gamma}} + N_0 \left(\frac{L_t}{\tau} \right)^{\frac{\gamma}{1-\gamma}} \right) \tag{21}$$

Agent j maximizes (20) over L_j subject to (21). Appendix B proves that the optimal learning decision of a representative agent is

$$L \equiv L_j = \max\{ \frac{A}{\frac{1-\gamma}{\gamma} + A} \overline{L}, \underline{L} \}, \forall j \in G_t, \forall t \tag{22}$$

where

$$A \equiv \frac{N_0 \tau^{-\frac{\gamma}{1-\gamma}}}{M + N_0 \tau^{-\frac{\gamma}{1-\gamma}}} + \frac{N_0}{M} \tau^{-\frac{1}{1-\gamma}} \tag{23}$$

By setting $N_0 = 0$ or $\tau = \infty$ which both correspond to the isolated families case (see section 4.1), we can verify that agents choose the lowest possible level of education $\underline{L}$. Also note that having no children (that is, choosing $L_j = \overline{L}$) is not possible in equilibrium: if none of j's neighbors has children, then j has no one to influence except his own children, hence a marginal decrease in the level of education and increase in the family size from zero to a small positive number would increase the influence/utility of j.

Note that the constant A consists of two components, both of which are positively related to N_0. The first component reflects the fact that when the network size N_0 increases, one's children are affected by an increasing number of non-parents. To retain influence on his own children, one has to study more. The second component is due to the fact that as the network size N_0 increases, one gets an opportunity to influence an increasing number of someone else's children. To maximize one's utility, it becomes optimal to reduce the number of own children and increase the quality of own idea.

We can also compute the optimal family size:

$$n = \min\{ \frac{\overline{L}}{\nu} \frac{1}{1 + \frac{\gamma}{1-\gamma} A}, \frac{1}{\nu} (\overline{L} - \underline{L}) \} \tag{24}$$

which decreases with an increase in the network size N_0. In the modern world, families are smaller, for the two reasons outlined above.

Another parameter of interest is the growth rate of the quality of ideas. From (8) and (14),

$$\frac{q_{t+1}}{q_t} = \frac{E_{t+1}^{\frac{1-\gamma}{\gamma}} L}{E_t^{\frac{1-\gamma}{\gamma}} L} = \left(\frac{E_{t+1}}{E_t}\right)^{\frac{1-\gamma}{\gamma}} = L\left(M + N_0 \tau^{-\frac{\gamma}{1-\gamma}}\right)^{\frac{1-\gamma}{\gamma}} \tag{25}$$

As the size of the network N_0 expands, and the rate of knowledge growth accelerates for two reasons. First, with access to knowledge of a wider network of predecessors, agents are able to learn more given the same learning time.[3] Second, a wider network of competitors induces people to spend more time on learning ($\frac{dL}{dN_0} > 0$), which also accelerates the rate of improvement of ideas.

4.3 Endogenous network size

Previous section assumes that the number of people from which one can learn is fixed. Now suppose that an agent can learn from *all* members of the previous generation, that is, the network size N_0 for $j \in G_t$ is equal to $N_{t-1} \equiv |G_{t-1}|$. With the increase of literacy rates, knowledge of foreign languages, and development of modern information technology, this assumption becomes increasingly realistic. As directly follows from this assumption, the network size now varies with the population size, and therefore the constant A defined in (23) is no longer a constant: it increases with the increase in population; denote by A_t the value of A for generation t. Formula n describes the family size, which decreases with A. Therefore, when the world population is small, the value of A_t is low, families are large, and the rate of population growth is high. As the world population gets larger, family size decreases until the size of the world population eventually arrives to a steady state. Given that n_t decreases with A_t, while A_t increases with population size, such a steady state is unique. This result is consistent with the forecast of the United Nations that population size will stabilize by the end of the 21-st century.

The rate of knowledge growth, vice versa, will accelerate as population rises. This result is similar to Michael Kremer's work on population growth and technological change (Kremer 1993), but the mechanism of acceleration is different. According to Kremer, knowledge grows faster in larger populations because of non-rivalry of knowledge. Assuming that a new invention can be utilized by everyone and assuming a constant number of innovators per capita, larger population means larger total number of innovators which means faster growth of technology. In my model, larger population means that each person has more people to influence and thus will accumulate more knowledge in order to look more appealing and to stay on par with his/her contemporaries who also acquire more knowledge.

[3]Mathematically, this effect is identical to the "love of variety" property of the Dixit-Stiglitz preferences widely used in International Trade literature

5 The effect of outside influence

People living in the province, especially in rural areas, have a very limited ability to influence the rest of the world; in this respect, they are not very much different from traditional societies that dominated the world two centuries ago. Nevertheless, demographic transition has occurred in the province too.

I propose the following explanation of the phenomenon. Agents choose to have fewer children not only when they face an opportunity to influence non-children, but also when their children can be influenced by the outside world. In traditional societies, parents are monopolists in the "ideological market" of their children, and hence have no incentive to improve the quality of their ideas. In the modern world, children can learn not only from their parents, but also from many other sources – travelers visiting their community, radio, television, and finally the internet. In this environment, a parent of fifteen children, who has spent all his life on raising children and thus has little education, will have very little influence even on his own children: they would choose to learn from television instead. A rational parent should reduce the family size and acquire more education, in order to be able to compete with the ideas delivered by the outside world.

In this section, I develop an extension of the model that formalizes the above intuition. I assume that at the beginning of their lives, all people are randomly divided into two groups – *artists* (fraction α of population) and *peasants* (fraction $1-\alpha$). Formally, the set G_t of all agents of generation t is partitioned into the subset of artists G_t^A and peasants G_t^P. As before, all agents are distributed uniformly on a circle. Given randomness of division into artists and peasants, agents within each of these groups are also uniformly distributed on that circle. Each agent $j \in G_t$, either artist or peasant, can learn from two sources. First, he can learn from own parent $\mathcal{C}^{-1}(j)$, with the learning cost $\tau_{\mathcal{C}}^{-1}(j)j = 1$ as before. Second, he can learn from the mass N_0 of the nearest artists; denote the set of these artists by $T(j)$. The cost of learning from artists is $\tau_{ij} = \tau \geq 1, i \in T(j)$. Learning from artists other than those belonging to $T(j)$, as well as from peasants other than one's parent, is impossible. Note that when $\alpha = 0$, we are back to the isolated families scenario analyzed in section (4.1).

I make an additional assumption that artists have no children and spend all their life developing their ideas. This assumption is motivated by the following several arguments. First, the main goal of this section is to analyze the effect of artists' influence on the fertility decisions of peasants; fertility decisions of artists themselves are of secondary interest. Second, when α is small enough, each artist has access to the minds of a large number of people, and having no children is an incentive compatible strategy. Third, when artists have no families, everyone's parent must be a peasant; therefore, any agent $j \in G_t$, whether artist or peasant, has the same learning environment $E_j = E_t$. As discussed above, symmetry of the learning environment simplifies the analysis by making one's learning decisions independent of his learning environment. Lastly, it is an empirical fact that many celebrities choose not to have children until their late-thirties, and spend little of their personal time on their nurture thereafter.

We now proceed to the analysis of learning and fertility decisions made by peasants. From (10), the learning environment faced by an agent $j \in G_t$, either artist or peasant, is

$$E_j = ME_{\mathcal{C}^{-1}(j)} \left(L_{\mathcal{C}^{-1}(j)}\right)^{\frac{\gamma}{1-\gamma}} + \int_{i \in T(j)} E_i \left(\frac{L_i}{\tau}\right)^{\frac{\gamma}{1-\gamma}} di$$

Note that learning environments of the predecessors are symmetric: $E_{\mathcal{C}^{-1}(j)} = E_i \equiv E_{t-1}$. Also, the learning choices of the artists are also symmetric and equal to their upper bound: $L_i = \overline{L}$. These considerations allow us to rewrite the above expression as follows:

$$E_j = E_{t-1} \left(M \left(L_{\mathcal{C}^{-1}(j)}\right)^{\frac{\gamma}{1-\gamma}} + N_0 \left(\frac{\overline{L}}{\tau}\right)^{\frac{\gamma}{1-\gamma}} \right) \tag{26}$$

We now adapt the expression for utility (11) for the environment defined in this section. The utility of artists is immaterial, since their choice of learning time is assumed rather than computed. The utility of peasants $j \in G_t^P$ is

$$U_j = L_j^{\frac{\gamma}{1-\gamma}} E_j M \frac{\overline{L} - L_j}{\nu} \frac{L_{\mathcal{C}(j)}}{E_{\mathcal{C}(j)}}$$

Note that peasants can influence only their own children.

From Proposition 2, children's learning choices $L_{\mathcal{C}(j)} = L_{t+1}$ do not depend on parent's decisions and can be treated by parents as given. The utility can be rewritten as follows:

$$U_j = L_{t+1} \frac{\overline{L} - L_j}{\nu} \frac{ML_j^{\frac{\gamma}{1-\gamma}}}{ML_j^{\frac{\gamma}{1-\gamma}} + N_0 \left(\frac{\overline{L}}{\tau}\right)^{\frac{\gamma}{1-\gamma}}} \tag{27}$$

If learning from artists is not possible ($N_0 = 0$ or $\tau = \infty$), parents do not have any competition with the outside world for the minds of their children and are best-off choosing minimal education and maximal family size. In a more general case, a closed-form solution for the optimal learning time does not exist; appendix C proves that increased access to artists' ideas induces peasants to learn more and have smaller families: $\frac{dL_j}{dN_0} > 0$, $\frac{dn_j}{dN_0} < 0$. Increased influence from the outside world induces parents to invest more in their own education, in order to be able to influence their own children.

6 Conclusion

This paper develops an "influential" theory of demographic transition, according to which the objective of agents in generation t is to influence the minds of the next generation. In traditional societies with poor communication technologies, it is difficult to influence anyone other than own children; in addition, children are unlikely to be influenced by anyone other than own parents. Both of these considerations induce parents to maximize the size of families and minimize educational effort. In modern societies with easy long-distance communication, competition for influence intensifies, which induces agents to spend more time on development of own ideas at the cost of smaller families.

The following extensions of this research may be of interest. First, a more general asymmetric learning environment (i.e. the environment in which agents have asymmetric access to the ideas of predecessors) can be studied. The difficulty of the analysis arises from the fact that in the asymmetric setting, the learning environment E_j of an agent $j \in G_t$ does not cancel out from his utility function and affects his optimal choice of learning time L_j. Since j's learning environment depends on educational decision of j's parent $L_{\mathcal{C}^{-1}(j)}$, we conclude that optimal L_j is a function of $L_{\mathcal{C}^{-1}(j)}$. Extrapolating this conclusion one generation forward, we find that the optimal learning decision of j's children, $L_{\mathcal{C}(j)}$, is an unknown function of j's own decision L_j. Therefore, computation of the optimal L_j (and therefore optimal family size) requires solving a differential equation, and most likely will not result in a closed-form solution.

Another potential extension is a generalization of the utility function. Since ideas of generation $t-1$ are embedded in the ideas of generation t, the influence of the latter on generation $t+1$ is also an indirect influence of $t-1$ on $t+1$. This indirect influence can be made part of the objective function of the members of generation $t-1$. This extension may help to explain why a parent $i \in G_{t-1}$ may choose to send her child $\mathcal{C}(i) \in G_t$ to learn from other (more educated) people $i' \in G_{t-1}$ at the universities. By doing so, i reduces the share of her own ideas in $\mathcal{C}(i)$, but increases the influence of $\mathcal{C}(i)$ on generation $t+1$; on balance, i may be better-off. For example, parents of Rudolph Diesel might have increased their own social influence by making their son learn from smart teachers and thus making him more influential.

Another possible extension of the model is modification of the production function of ideas (2). Instead of assuming that the elasticity of substitution between any two ideas is constant, one could assume that all ideas (and thus all people) are divided into two sets, or two *religions*. Elasticity of substitution between two ideas from different religions is infinite, and therefore an individual can choose to learn from representatives of only one religion. Given that parents enter the learning environment of their children with a strictly positive weight, the latter are naturally biased towards the religion of their parents. But if the "other" religion is sufficiently influential, children may choose to switch to that other religion, dramatically reducing the influence (utility) of their parents. In this environment, one could formalize a theory that celibacy of Christian priests has helped them increase their influence in the society and convert a large part of the world into Christianity (see (Blackmore 2000) for a verbal description of the theory).

A Derivation of demand for ideas

Maximization of (2) subject to (3) over all learning intensities x_{ij} is equivalent to the following unconstrained maximization:

$$\max_{x_{ij},\forall i\in G_{t-1}}\left\{q_j-\lambda\left[L_j-Mx_{\mathcal{C}^{-1}(j)j}\tau_{\mathcal{C}^{-1}(j)j}-\int_{i\in G_{t-1}\backslash\mathcal{C}^{-1}(j)}x_{ij}\tau_{ij}di\right]\right\} \tag{28}$$

where q_j is defined in (2), and λ is the Lagrange multiplier. Maximization with respect to an arbitrary x_{ij} yields

$$q_j^{\frac{1-\gamma}{\gamma}}x_{ij}^{-(1-\gamma)}q_i^{\gamma}-\lambda\tau_{ij}=0 \tag{29}$$

Note that this expression applies to both learning from a parent $i=\mathcal{C}^{-1}(j)$ and a non-parent $i\neq\mathcal{C}^{-1}(j)$. We can now solve for x_{ij}:

$$x_{ij}=C_jq_i^{\frac{\gamma}{1-\gamma}}\tau_{ij}^{-\frac{1}{1-\gamma}} \tag{30}$$

where C_j is some positive constant. We can now substitute (30) into (3), keeping in mind (4), to obtain

$$C_jq_{\mathcal{C}_j^{-1}}^{\frac{\gamma}{1-\gamma}}+\int_{i\in G_{t-1}\backslash\mathcal{C}^{-1}(j)}C_j\left(\frac{\gamma}{1-\gamma}\right)^{\frac{\gamma}{1-\gamma}}di=L_j \tag{31}$$

By using the definition of the learning environment (6), we can simplify the above formula to simply $C_jE_j=L_j$, or $C_j=\frac{L_j}{E_j}$. By substituting the latter expression into (30), we end up with the optimal learning intensities presented in (7).

B Learning from parent's neighbors: derivation of optimal learning time

In the expression for utility (20), the multiplier $L_{t+1}E_t$ is positive and constant and thus can be ignored when searching for argmaximum. The first order condition is then as follows:[4]

$$\frac{\gamma}{1-\gamma}L_j^{\frac{\gamma}{1-\gamma}-1}\left[M\frac{\overline{L}-L_j}{\nu}\frac{1}{E_{\mathcal{C}(j)}}+n_tN_0\frac{1}{E_{t+1}}\tau^{-\frac{1}{1-\gamma}}\right]-L_j^{\frac{\gamma}{1-\gamma}}M\frac{1}{\nu}\frac{1}{E_{\mathcal{C}(j)}}-L_j^{\frac{\gamma}{1-\gamma}}M\frac{\overline{L}-L_j}{\nu}\frac{1}{E^2_{\mathcal{C}(j)}}\frac{dE_{\mathcal{C}(j)}}{dL_j}=0 \tag{32}$$

[4]To simplify exposition, we abstract from upper and lower bounds on L_j; the solution (22) presented in the main body of the paper does account for the bounds.

In symmetric equilibrium, $E_{\mathcal{C}(j)} = E_{t+1}$ which now can be factored out; $L_j = L_t$. By dividing the above expression by $L_j^{\frac{\gamma}{1-\gamma}-1}\frac{1}{E_{t+1}}$, we obtain

$$\frac{\gamma}{1-\gamma}\left[M\frac{\overline{L}-L_t}{\nu}+n_t N_0\tau^{-\frac{1}{1-\gamma}}\right]-L_t M\frac{1}{\nu}-L_t M\frac{\overline{L}-L_t}{\nu}\frac{1}{E_{t+1}}\frac{dE_{\mathcal{C}(j)}}{dL_j}=0 \tag{33}$$

Here, the last component can be simplified to

$$\begin{aligned}
&L_t\,M\frac{\overline{L}-L_t}{\nu}\frac{1}{E_{t+1}}\frac{dE_{\mathcal{C}(j)}}{dL_j}\\
=\;&L_t\,M\frac{\overline{L}-L_t}{\nu}\frac{1}{E_t L_t^{\frac{\gamma}{1-\gamma}}\left(M+N_0\tau^{-\frac{\gamma}{1-\gamma}}\right)}E_t\frac{\gamma}{1-\gamma}ML_t^{\frac{\gamma}{1-\gamma}-1}\\
=\;&M\frac{\overline{L}-L_t}{\nu}\frac{\gamma}{1-\gamma}\frac{M}{M+N_0\tau^{-\frac{\gamma}{1-\gamma}}}
\end{aligned} \tag{34}$$

Substituting (34) back into (33) and substituting the formula for optimal family size $n_t = \frac{\overline{L}-L_t}{\nu}$, we obtain

$$\frac{\gamma}{1-\gamma}\frac{\overline{L}-L_t}{\nu}\left[M+N_0\tau^{-\frac{1}{1-\gamma}}\right]-L_t M\frac{1}{\nu}-M\frac{\overline{L}-L_t}{\nu}\frac{\gamma}{1-\gamma}\frac{M}{M+N_0\tau^{-\frac{\gamma}{1-\gamma}}}=0 \tag{35}$$

After some manipulations, we obtain

$$\frac{\gamma}{1-\gamma}\frac{\overline{L}-L_t}{\nu}\left[M\frac{N_0\tau^{-\frac{\gamma}{1-\gamma}}}{M+N_0\tau^{-\frac{\gamma}{1-\gamma}}}-N_0\tau^{-\frac{1}{1-\gamma}}\right]=L_t M\frac{1}{\nu} \tag{36}$$

By dividing both sides by $\frac{M}{\nu}$, we get

$$\frac{\gamma}{1-\gamma}\left(\overline{L}-L_t\right)A=L_t \tag{37}$$

where A is defined in (23). Trivial manipulations with the above formula yield the optimal learning time (22).

C Outside influence: derivation of optimal learning time

Define

$$G(x,y)\equiv\frac{Mx^{\frac{\gamma}{1-\gamma}}}{Mx^{\frac{\gamma}{1-\gamma}}+y\left(\frac{\overline{L}}{\tau}\right)^{\frac{\gamma}{1-\gamma}}}$$

Then, logarithm of utility (27) is

$$F(L_j, N_0) \equiv \log U_j(L_j, N_0) = \log \frac{L_{t+1}}{\nu} + \log(\overline{L} - L_j) + \log G(L_j, N_0) \tag{38}$$

Abstracting from the bounds on L_j, the first-order condition for the optimal learning time is

$$\frac{dF}{dL_j} = -\frac{1}{\overline{L} - L_j} + \frac{\frac{dG}{dL_j}(L_j, N_0)}{G(L_j, N_0)} = 0 \tag{39}$$

By computing $\frac{dG}{dL_j}$, we can show that

$$\frac{dF}{dL_j} = -\frac{1}{\overline{L} - L_j} + \frac{\gamma}{1-\gamma}\frac{1}{L_j}(1 - G(L_j, N_0)) \tag{40}$$

From the implicit function theorem,

$$\frac{dL_j}{dN_0} = -\frac{\frac{d^2F}{dL_j dN_0}}{\frac{d^2F}{dL_j^2}} \tag{41}$$

It is straightforward to verify that $\frac{d^2F}{dL_j dN_0} > 0$ and $\frac{d^2F}{dL_j^2} < 0$; therefore, $\frac{dL_j}{dN_0} > 0$.

References

Axinn, William G. and Jennifer S. Barber, "Mass Education and Fertility Transition," *American Sociological Review*, 2001, *66* (4), 481–505.

—— **and Scott T. Yabiku**, "Social Change, the Social Organization of Families, and Fertility Limitation," *The American Journal of Sociology*, 2001, *106* (5), 1219–1261.

Becker, Gary, "An economic analysis of fertility," in "Demographic and Economic Change in Developed Countries," Princeton: National Bureau of Economic Research, 1960, pp. 209–231.

Becker, Gary S., Kevin M. Murphy, and Robert Tamura, "Human Capital, Fertility, and Economic Growth," *The Journal of Political Economy*, 1990, *98* (5), S12–S37.

Behrman, Jere R., Hans-Peter Kohler, and Susan Cotts Watkins, "Social Networks and Changes in Contraceptive Use over Time: Evidence from a Longitudinal Study in Rural Kenya," *Demography*, 2002, *39* (4), 713–738.

Blackmore, Susan, *The meme machine*, Oxford University Press, 2000.

Bongaarts, John and Susan Cotts Watkins, "Social Interactions and Contemporary Fertility Transitions," *Population and Development Review*, 1996, *22* (4), 639–682.

Boyd, Robert and Peter J. Richerson, *Culture and the evolutionary process*, University of Chicago Press, 1985.

Caldwell, John C., "Toward A Restatement of Demographic Transition Theory.," *Population and Development Review*, 1976, *2* (3/4), 321 – 366.

Cosmides, Leda and John Tooby, "Better than Rational: Evolutionary Psychology and the Invisible Hand," *The American Economic Review*, 1994, *84* (2), 327–332.

Dawkins, Richard, *The Selfish Gene*, third ed., Oxford University Press, 2006 (1976).

Galor, Oded and Omer Moav, "Natural Selection and the Origin of Economic Growth," *The Quarterly Journal of Economics*, 2002, *117* (4), 1133–1191.

Greksa, L. P., "Population growth and fertility patterns in an Old Order Amish settlement.," *Annals of Human Biology*, 2002, *29* (2), 192 – 201.

Henrich, Joseph and Francisco J. Gil-White, "The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission," *Evolution and Human Behavior*, 2001, *22* (3), 165–196.

Kremer, Michael, "Population Growth and Technological Change: One Million B.C. to 1990," *The Quarterly Journal of Economics*, 1993, *108* (3), 681–716.

Newson, Lesley, Tom Postmes, S. E. G. Lea, and Paul Webley, "Why Are Modern Families Small? Toward an Evolutionary and Cultural Explanation for the Demographic Transition," *Personality and Social Psychology Review*, 2005, *9* (4), 360–375.

Rae, John, *Life of Adam Smith*, ninth ed., Cosimo, Inc, 2006 (1895).

Tooby, John and Leda Cosmides, "The Psychological Foundations of Culture," in Jerome H. Barkow, Leda Cosmides, and John Tooby, eds., *The adapted mind: evolutionary psychology and the generation of culture*, Oxford University Press, 1992, pp. 19–136.