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

Zakharenko, Roman

State University - Higher School of Economics (Moscow)

14 April 2010

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

Roman Zakharenko*

April 14, 2010

Abstract

In this paper, I propose a theory of demographic transition and associated growth of the stock of knowledge that posits *social influence* as a key determinant of these phenomena. Individuals are influenced by *ideas* developed in the previous generation; their goal is to maximize their own social influence, that is, the extent to which their ideas have been used by the next generation. 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, Z13

*International College of Economics and Finance, State University Higher School of Economics, Moscow, Russia. Email: rzakharenko@hse.ru. Web: www.rzak.ru. I thank Maxim Arnold, Kenneth Binmore, Ronald Lee, Lesley Newson, my colleagues at ICEF, seminar participants at SUHSE, and participants of the Midwest Economic Theory conference, who's ideas have influenced this research.

1 Introduction

The idea that one tries to produce as many offspring as possible, and responds to increase in resources by increase in fertility, belongs to Thomas Malthus, and was originally applied only to humans. Only half a century later did Darwin generalize this approach to all living organisms, from bacteria to primates. Ironically, today this theory appears to be valid for all life on Earth *except* humans for whom it was originally designed. In the past two hundred years, the world underwent a profound demographic transition, and the number of surviving children produced by humans in most societies is far below their "production possibility frontier". The goal of this paper is to offer a theory of demographic transition, building on recent developments in fields adjacent to Economics, and sticking to the (broadly defined) Darwinian philosophy. The paragraphs that follow outline 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 with respect to their quality 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.

The next section discusses the biological foundations of influence-based (rather than consumption-based) preferences.

1.1 Influence-based preferences

Beginning about three decades ago, theoretical biologists and anthropologists have outlined one key difference of the *Homo sapiens* from all other living species – their unprecedented ability to imitate each other's behavior. This ability gives rise to *culture*, which Boyd and Richerson [2005] (page 3) define as follows.

> Culture is information that people acquire from others by teaching, imitation, and other forms of social learning. On a scale unknown in any other species, people acquire skills, beliefs, and values from the people around them, and these beliefs strongly affect behavior.

It is important that unlike transmission of genes that is always vertical (that is, goes from parent to child), transmission of culture may be not only vertical but also horizontal, i.e. from peer to peer. Boyd and Richerson [2005] (Chapter 1.1) argue that the evolutionary purpose of cultural transmission is that it serves as a "shortcut" for the learners: it allows them to tremendously accelerate the speed of adaptation to a rapidly changing environment by imitating the most successful individuals they observe. Henrich and Gil-White [2001] argue that those who imitate may develop a psychological attachment to those being imitated (the "role models"), which makes the latter more *prestigious*, or gives them higher social status. But in the hunter-gatherer society in which the human psychology has evolved, I argue, higher social status has been associated with higher genetic fitness, and hence individuals might have evolved a preference for a higher social influence.

Modern society, however, is very different in many ways from the hunter-gatherer society for which human psychology has been designed. In particular, unlike hunter-gatherer societies in which humans were influenced mainly by parents and close relatives, modern humans routinely interact with and are influenced by hundreds or thousands of strangers; they are also influenced by scientists whose ideas they learn at school, and celebrities they see on television. In other words, with lower costs of communication, increased mobility of people, and increased urbanization rates, competition for social influence has tremendously intensified. In this environment, I argue, the preference for social influence may go against the original purpose of increasing genetic fitness: being influential in the modern environment means spending a lot of time on development of one's skills, which leaves less time for family and children.

The literature in Economics on cultural transmission originates in the late 1990s with the works of Bisin and Verdier [1998] and Bisin and Verdier [2001]. Like the current paper, they emphasize the dichotomy of vertical (from parents to children) and horizontal, or oblique (between non-relatives) transmission, and also assume that parents derive utility from their cultural traits being vertically transmitted to children. Bisin et al. [2004] find empirical evidence of religious parents undertaking costly actions to influence their children to adopt the same religion.

The assumption made in this paper that individuals are also willing to influence non-relatives is new, but, as elaborated above, it can be related to large literature on people caring about own social status, since status goes hand-in-hand with influence. In the formulation of Henrich and Gil-White [2001], high-status (or prestigious as they label them) individuals are *by definition* those who influence the behavior of others. The examples of research on status-based preferences and their implications include Cole et al. [1992] and Robson [1992].

Unlike the existing literature in Economics that assumes private-consumption-based utility or, at best, mixes private consumption with concern for status and/or altruism towards children, I entirely exclude consumption from individual utility and assume that utility depends on exerted influence only. In theoretical biology, consumption of food is not viewed as a final good entering utility, but as an investment into bodily/somatic capital that enables one to pursue the ultimate goal of maximizing one's genetic fitness in the future. In human society, additionally, part of consumption can be viewed as increasing one's social status/influence; inclusion of consumption into the utility function thus has no theoretical justification.[1] Even Adam Smith, the inspirer of the "Homo economicus" concept, in his *Theory of Moral Sentiments* famously argued:

> To what purpose is all the toil and bustle of the world?... It is not wealth that men desire, but the consideration and good opinion that wait upon riches.

Moreover, the desire for higher consumption is not universal as in many societies modesty and abstention are viewed to be more appropriate, whereas higher social influence and greater respect are universally desired.

[1] Robson [2001] justifies inclusion of personal consumption into utility by a bounded rationality argument: if the amount of consumption is highly correlated with the "true" utility function, but optimization of the latter is computationally intensive, individuals can maximize a consumption-based utility instead.

1.2 Related theories of demographic transition

1.2.1 "Influential" theories

The theory of demographic transition developed in this paper is most closely related to that of Newson et al. [2007] as it identifies the same phenomenon – increased interaction with non-relatives – as the key driver of demographic transition. The mechanism developed in Newson et al. [2007], however, is different. Individuals are influenced by close relatives who encourage them to have more children (the "inclusive fitness" argument, Hamilton [1964]). In traditional societies, influence from relatives is large and hence fertility is high, while in modern societies there is no such influence, which results in low fertility. In contrast, the theory I propose explains low fertility not by lack of influence from relatives, but by presence of influence from non-relatives. To distinguish between the two theories, one should measure fertility in families that live in isolation, i.e. exposed to influence of neither relatives nor non-relatives. If Newson et al. [2007] are right, lack of influence from relatives would result in modern small families; if the theory I propose is correct, lack of influence from non-relatives would leave families large. Anecdotal evidence suggests that in preindustrial societies, parents who lived in isolation or near isolation had far more children than modern parents. For example, the famous Lykov family, which retreated from the Communist oppression into the Siberian taiga in 1932 and lived for nearly fifty years in full isolation, had four children, three of which were born during the retreat (Peskov [1994]).

Another related theory of demographic transition is proposed by Blackmore [2000] (pages 139-142). The theory builds on the idea of *memes* originally proposed by Dawkins [1976]. The fact that humans can get influenced by each other's ideas gives rise to *memes*, or viruses of the mind, that alter human behavior so as to spread themselves into more and more brains. In traditional societies with isolated from each other communities, the spread of memes is limited, and the standard Darwinian desire to produce as many surviving children as possible dominates. In the modern world, low costs of communication amplify the role of memes, which induce their host humans to spend more time on spreading themselves, leaving less time for physical reproduction. This approach, however, was not accepted by the academic literature, due to impossibility to verify empirically the existence of the memes.

1.2.2 Socioeconomic theories

In the field of Economics, the existing literature associates the fertility decline with improving socioeconomic conditions of the households. The modern mainstream theories of demographic transition build on contributions of Becker [1960] and Becker and Lewis [1973], the main idea of which is that children should be viewed as a consumption good which requires a time input to enjoy; fertility declines in response to increased earnings opportunities of household members (mainly women), who switch from enjoying the number of children to enjoying their "quality", as well as consumption of other goods. A formal analysis that relates fertility and income growth was pioneered by Becker et al. [1990] who develop a multiple equlibria model, in which low-income equilibrium is characterized by low return to human capital which results in high fertility and zero per-capita income growth, while a high-income equilibrium is associated with high return to human capital, low fertility, and rapid growth. More recent contributions in this stream of literature include Galor and Weil [1996] who argue that technological progress, which has diminished the role of muscle power and increased the role of brainpower, increased the economic opportunities of women and thus made them substitute children for consumption; Galor and Weil [2000] build a more accurate model of demographic transition in which fertility first rises and then declines in response to income growth, which is consistent with the European historic trend.

Another popular in Economics explanation of declining fertility is the "old age security hypothesis" by Caldwell [1976] who argues that when social security is poor, parents have many children, expecting that children will support them in their old age; improved social security reduces the need for children and therefore results in a lower fertility.

Finally, a third popular stream of literature attributes declining fertility to a declining child mortality rate: if parents are risk averse, they respond to better health services by sharply reducing the number of births, which results in a smaller number of surviving children. Recent examples of this school of thought include Kalemli-Ozcan [2003] and Tamura [2006].

It seems very plausible that all above mentioned factors of fertility decline – increased return to human capital (especially that of women), better social security, lower child mortality – are often associated with increased interaction of individuals with non-relatives, and thus to increased exposure to outside cultural influence. Therefore, the three above factors may be correlated with unobserved social influence, rather than affect fertility *per se*. In section 5, I provide empirical evidence that at least part of fertility decline should be attributed to the effect of outside cultural influence, rather than to the effect of socioeconomic change.

2 Model

2.1 Basics

This is an overlapping-generations model, where each generation lives for two periods. Each generation is labeled by its period of birth, $t \in \{0, 1, \dots\}$. In each generation, there is a continuum G_t of individuals. It is assumed that individuals make all their decisions when young, and enjoy their utility when old. For simplicity, there is no time discounting.

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

An individual $j \in G_t$ is influenced by ("influence"is rigorously specified in section 2.2) his parent $C^{-1}(j) \in G_{t-1}$, as well as by a subset of parents' contemporaries, or "neighbors". To define formally the set of parent's neighbors, it is convenient to assume that there exists a spatial allocation of members of a given generation; I assume they are distributed uniformly on a circle. The density of agents on the circle is constant, so the circle expands with population growth. There is a metric that measures distance between any two members of a given generation. For the analysis that follows, for every $j \in G_t, \forall t$ I define a set $N(j)$ of neighbors of j's parent as the set of all $i \in G_{t-1} \backslash C^{-1}(j)$ such that $|N(j)| = N$ and $||i, C^{-1}(j)|| \leq ||i', C^{-1}(j)||$ for any $i \in N(j)$ and $i' \in G_{t-1} \backslash N(j) \backslash C^{-1}(j)$. I assume that total population is large enough, so that $N \leq |G_t|$ for every time period t. I refer to members of $N(j)$ as *non-parents* of j.

It is also convenient to define the set of individuals in the next generation who are influenced by $j \in G_t$. By the *influence set* $S(j)$ we will denote all such individuals, except j's own children $C(j)$. Formally, $S(j)$ is the set of all $k \in G_{t+1}$ such that $j \in N(k)$. Figure 1 illustrates the geography of influence.

The rest of the model is presented in a somewhat unusual order: I first define individual constraints and then the utility function.

2.2 Constraints

Each individual $j \in G_t, \forall t$ is endowed with L units of time that can be utilized in two activities: learning 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 and spends $x(j)$ units of time

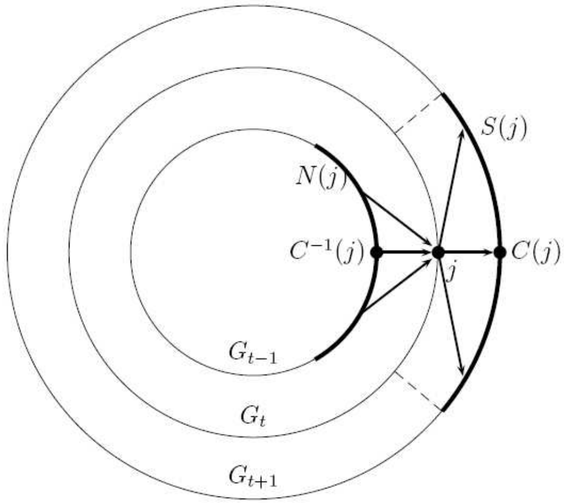


Figure 1: An illustration of the geography of influence

learning has the following budget constraint:

$$x(j) + \nu n(j) \leq L \tag{1}$$

Individuals are allowed to choose non-integer number of children, which can be interpreted as follows. While the choice of a parent $n(j)$ 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.

By assumption, there exists a learning lower bound: $x(j) \geq \underline{L} > 0$. This assumption reflects the fact early in their life, individuals can only learn and cannot reproduce.

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, and which can be viewed as an analog of human capital in conventional models of demographic transition. To develop an idea, j has to absorb, or get *influenced* by, 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, $C^{-1}(j)$, has a strictly positive and finite influence on the formation of j's idea, while each of non-parents $N(j) \subset G_{t-1}$ has an infinitesimal influence on j. The aggregate influence of non-parents is of the same order of magnitude as the influence of the parent. Formally,

$$q(j) = \left(\left[q(C^{-1}(j)) y(C^{-1}(j), j) \right]^{\frac{\sigma-1}{\sigma}} + \int_{i \in N(j)} \left[q(i) y(i,j) \right]^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}} \tag{2}$$

Here $y(i,j) \geq 0$ is the influence of i on j; $\sigma > 1$ is the elasticity of substitution between ideas.[2] While $q(i)$ is given for j, the extent to which he is influenced by i, $y(i,j)$, is his control parameter, and is chosen subject to a time constraint outlined below.

Absorbtion of ideas incurs a time cost per unit of influence. I assume that the time cost of learning from one's parent is normalized to unity, while the cost of learning from non-parents is $\tau \geq 1$. The latter is the key parameter of interest in this paper: falling costs of learning from non-parents imply increased outside influence, and affect fertility and accumulation of knowledge.

The total amount of time spent on learning from all sources must add up to $x(j)$:

$$y(C^{-1}(j), j) + \int_{i \in N(j)} y(i,j) \tau di = x(j) \tag{3}$$

2.3 Objective function

Individuals derive utility from increasing the share of their ideas in the minds of the next generation; they place a special emphasis on their own children. The extent to which j's child has absorbed his/her ideas has a finite positive weight on j's utility, while the influence on each member of $S(j)$ has an infinitesimal weight. The influence on children and the aggregate influence on those in $S(j)$ are of the same order of magnitude. Formally,

$$U(j) = \sum_{k \in C(j)} \frac{y(j,k)}{y(j,k) + \int_{j' \in N(k)} y(j',k) dj'} + \int_{k \in S(j)} \frac{y(j,k)}{y(C^{-1}(k), k) + \int_{j' \in N(k)} y(j',k) dj'} dk \tag{4}$$

[2]The constant elasticity of substitution function of ideas production was chosen to reflect the fact that individuals are influenced more by better ideas, and respond smoothly to changes in the quality of ideas.

2.4 Initial conditions

It is assumed that every individual in the initial generation G_0 has n_0 children and the quality of idea of q_0.

3 Analysis

3.1 Optimal learning intensities

There are many possible equilibria in the environment described above; I concentrate on the "focal" equilibria in which individuals are influenced more by better ideas. The Proposition below shows that this property is recursive: if it is true for the influence of generation t on generation $t+1$, it is also true for the influence of $t-1$ on t.

For the analysis that follows, it is convenient to define the following variable:

$$E(j) \equiv q(C^{-1}(j))^{\sigma-1} + \int_{i \in N(j)} \left(\frac{q(i)}{\tau} \right)^{\sigma-1} di \tag{5}$$

$E(j)$ is 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.

Proposition 1 *If a better idea $q(j)$ of $j \in G_t$ increases his influence on the next generation, the optimal extent to which he is influenced by $i \in \{C^{-1}(j) \cup N(j)\}$ is*

$$y(i,j) = \frac{x(j)}{E(j)} q(i)^{\sigma-1} \tau^{-\sigma \times I(j \notin C(i))} \tag{6}$$

where $I(\cdot)$ is the indicator function.

The influence of i on j increases linearly with the overall length of j's education $x(j)$; it increases with the quality of i's idea $q(i)$, and decreases with learning cost τ, if i is a non-parent. The influence of i also decreases as j's learning environment $E(j)$ improves.

The intuition of the proof is as follows. If the influence of j on the next generation increases with the quality of his idea $q(j)$, he is willing to attain the highest possible $q(j)$ for any given value of his learning time $x(j)$. Therefore, the extent to which j is influenced by each member of the previous generation can be found by maximizing (2) over $y(i,j), \forall i \in$

$\{C^{-1}(j) \cup N(j)\}$ subject to (3). Appendix A shows that the optimal solution to this problem is given by (6).

From (2) and (6), we can also compute the optimal (highest possible) quality of j's idea given his learning time $x(j)$:

$$\begin{aligned} q(j) &= \frac{x(j)}{E(j)} \left(q(C^{-1}(j))^{\sigma-1} + \int_{i \in N(j)} \left(\frac{q(i)}{\tau} \right)^{\sigma-1} di \right)^{\frac{\sigma}{\sigma-1}} = \frac{x(j)}{E(j)} E(j)^{\frac{\sigma}{\sigma-1}} \\ &= E(j)^{\frac{1}{\sigma-1}} x(j) \end{aligned} \tag{7}$$

Using (7), we can rewrite (6) and (5):

$$y(i,j) = \frac{x(j)}{E(j)} E(i) x(i)^{\sigma-1} \tau^{-\sigma \times I(j \notin C(i))} \tag{8}$$

$$E(j) = E(C^{-1}(j)) x(C^{-1}(j))^{\sigma-1} + \int_{i \in N(j)} E(i) \left(\frac{x(i)}{\tau} \right)^{\sigma-1} di \tag{9}$$

Using (8) extrapolated one generation forward and canceling out $\frac{x(k)}{E(k)}$ from the numerators and denominators of both components, we can rewrite the utility (4) as follows:

$$\begin{aligned} U(j) &= \sum_{k \in C(j)} \frac{E(j)x(j)^{\sigma-1}}{E(j)x(j)^{\sigma-1} + \int_{j' \in N(k)} E(j')x(j')^{\sigma-1}\tau^{-\sigma} dj'} \\ &+ \int_{k \in S(j)} \frac{E(j)x(j)^{\sigma-1}\tau^{-\sigma}}{E(C^{-1}(k))x(C^{-1}(k))^{\sigma-1} + \int_{j' \in N(k)} E(j')x(j')^{\sigma-1}\tau^{-\sigma} dj'} dk \end{aligned} \tag{10}$$

The next section focuses on the study of the learning choice of a particular individual $j \in G_t$ (henceforth the decision maker), assuming that all other individuals within the same generation, as well as in other generations, make the same learning choices. To simplify the exposition of the analysis that follows, additional notations and assumptions are made.

- Since all members of all generations except $j \in G_t$ make identical learning choices, I omit subscripts and denote them by x.

- The learning environment and the quality of ideas are the same for all individuals within a generation but may change over time; I denote them by E_t and q_t, respectively.

- To emphasize that the equilibrium decisions depend on a measure of outside influence, $\overline{N} \equiv N\tau^{-\sigma}$, I display all relevant variables as functions of this parameter.

- By z I denote a learning choice of the decision maker, by $z(x, \overline{N})$ – his best response to the learning choices of others, given the amount of outside influence $\overline{N}$.

- From (1), the expected number of children of the decision maker is $\frac{L-z}{\nu}$. The expected number of children of his contemporaries is $\frac{L-x}{\nu}$. Given that each individual in the next generation is influenced by N non-parents, the measure of the influence set of the decision maker (as well as his contemporaries) is $|S(j)| = \frac{L-x}{\nu}N$.

With new notations, the objective function of the decision maker (10) can be rewritten as

$$U(z, x, \overline{N}) = \frac{L-z}{\nu} B(z, x, \overline{N}) + \frac{L-x}{\nu} H(z, x, \overline{N}) \tag{11}$$

where

$$B(z, x, \overline{N}) \equiv \frac{z^{\sigma-1}}{z^{\sigma-1} + N x^{\sigma-1} \tau^{-\sigma}} = \frac{z^{\sigma-1}}{z^{\sigma-1} + x^{\sigma-1}\overline{N}} \tag{12}$$

$$H(z, x, \overline{N}) \equiv \frac{N z^{\sigma-1} \tau^{-\sigma}}{x^{\sigma-1} + N x^{\sigma-1} \tau^{-\sigma}} = \frac{z^{\sigma-1}\overline{N}}{x^{\sigma-1}(1+\overline{N})} \tag{13}$$

Intuitively, $B(z, x, \overline{N})$ is the share of the decision maker's ideas in the minds of his own children, while $H(z, x, \overline{N})$ is a measure of the decision maker's influence on the minds of others. Maximization of (11) over $z \in [\underline{L}, L]$ yields the solution to the problem of the decision maker.

4 Results

4.1 A primer: isolated families

Consider an extreme case in which children can only learn from their parents, that is $\tau = \infty$ and $\overline{N} = 0$. Then, $B(z, x, \overline{N}) \equiv 1$ and $H(z, x, \overline{N}) \equiv 0$; maximization of (11) over z yields the corner solution: $z(x) = \underline{L}$. When families cannot 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 Symmetric influence

We now focus on a general case in which the outside influence is positive: $\overline{N} > 0$. In this case, the argminimum of utility is $z = 0$; the argmaximum must be strictly positive. For the clarity of exposition, I assume that the educational lower bound $\underline{L}$ is low enough and does not bind the decision maker, which allows us to ignore it in the analysis that follows. The first derivative of utility (11) over z is (omitting function arguments)

$$\begin{aligned}\nu U_z &= (L-z)B_z - B + (L-x)H_z = (L-z)\frac{\sigma-1}{z}B(1-B) - B + (L-x)\frac{\sigma-1}{z}H \\ &= BF + (L-x)\frac{\sigma-1}{z}H \qquad (14)\end{aligned}$$

where

$$F(z,x,\overline{N}) \equiv (\sigma-1)\left(\frac{L}{z}-1\right)(1-B(z,x,\overline{N})) - 1 \qquad (15)$$

At the optimal point $z(x,\overline{N})$, the value of (14) must be zero if $z(x,\overline{N}) < L$, and nonnegative if $z(x,\overline{N}) = L$.

In (14), the properties of its components are as follows. $B(z,x,\overline{N})$ monotonically increases from zero to one as z increases from zero to infinity; $F(z,x,\overline{N})$ monotonically decreases from infinity to -1 as z increases from zero to L; $H(z,x,\overline{N})$ is positive, increasing and convex in z, hence the second component of (14) is also positive and increasing in z. Denote by $z^*(x,\overline{N})$ the value of z that makes F equal to zero:

$$F(z^*(x,\overline{N}),x,\overline{N}) \equiv 0 \qquad (16)$$

For $z \in (0, z^*(x,\overline{N})]$, (14) is strictly positive, and therefore the argmaximum of utility must be searched on $(z^*(x,\overline{N}), L]$.

The following Lemma is helpful in the analysis that follows.

Lemma 1 1. *Function $U(z,x,\overline{N})$ has at most two local maxima on $z \in (0,L]$ for given $x \in (0,\infty)$ and $\overline{N} > 0$.*

2. *Function $U(z,x,\overline{N})$ has at most one local maximum on $z \in (0,L)$ for given $x \in (0,\infty)$ and $\overline{N} > 0$ (and therefore the other one, if exists, is at $z = L$).*

3. *If*

$$B(L,x,\overline{N}) \leq \frac{\sigma}{2(\sigma-1)} \qquad (17)$$

function $U(z, x, \overline{N})$ *has only one maximum on* $z \in (0, L]$*, for given* $x \in (0, \infty)$ *and* $\overline{N} > 0$*.*

Note that when $\sigma < 2$, the condition (17) holds for any values of x and $\overline{N}$, because $B(L, x, \overline{N}) \leq 1$ by definition of B.

The proof of Lemma 1 is contained in Appendix B. Lemma 1 entails the following

Corollary 1 *Any point* $z_1 \in (0, L)$ *of utility local maximum, for given* x *and* $\overline{N}$*, is also the point of global maximum if and only if* $U(z_1, x, \overline{N}) \geq U(L, x, \overline{N})$*.*

We now turn to the central result of this section – characterization of the symmetric equilibrium, in which the decision maker's learning best response is equal to the learning decision of others, that is $z(x, \overline{N}) = x$.

Proposition 2 *(Symmetric pure strategy equilibrium) The equilibrium learning decision is*

$$\overline{x}(\overline{N}) = L\frac{(\sigma - 1)A(\overline{N})}{1 + (\sigma - 1)A(\overline{N})} \tag{18}$$

where

$$A(\overline{N}) \equiv \frac{\overline{N}}{1 + \overline{N}} + \overline{N} \tag{19}$$

if and only if

$$(\sigma - 1)\log\left(1 + \frac{1}{A(\overline{N})(\sigma - 1)}\right) < \log\left(1 + \frac{1}{\overline{N}}\right) \tag{20}$$

An increasing exposure of individuals to outside influence $\overline{N}$, brought about by falling learning costs τ, increases $A(\overline{N})$, thereby increasing the learning time $\overline{x}(\overline{N})$ and decreasing the family size $n(\overline{N}) = \frac{L - \overline{x}(\overline{N})}{\nu}$.

Proof. The expression (18) for the steady state learning time $\overline{x}(\overline{N})$ was derived by setting $U_z(\overline{x}, \overline{x}, \overline{N})$ equal to zero and solving for $\overline{x}$, and therefore the first-order condition of maximum holds by construction of $\overline{x}$. The second derivative of utility can be shown to be

$$\nu U_{zz}(\overline{x}, \overline{x}, \overline{N}) = -\frac{\sigma - 1}{\overline{x}}\frac{\overline{N}}{A(\overline{N})(1 + \overline{N})^2}\left[\overline{N} + \frac{2 + \overline{N}}{\sigma - 1}\right] < 0 \tag{21}$$

therefore the second-order condition also holds. By Corollary 1, $\overline{x}(\overline{N})$ refers to the maximum of utility if $U(\overline{x}(\overline{N}), \overline{x}(\overline{N}), \overline{N}) \geq U(L, \overline{x}(\overline{N}), \overline{N})$, which is equivalent to (20). ■

It can be shown that the condition (20) holds for any $\overline{N}$ if $\sigma < 2$, and for any σ if $\overline{N} > \overline{N}_0 \approx \frac{3}{4}$. Therefore, symmetric pure strategy equilibrium does *not* exist only when

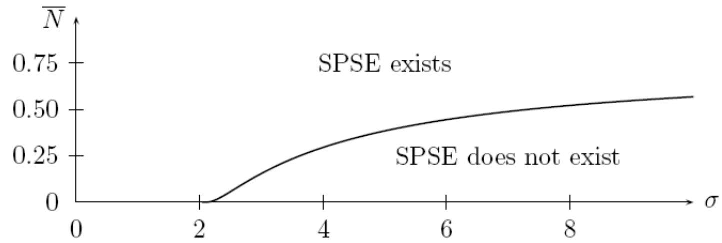


Figure 2: The region of existence of the symmetric pure strategy equilibrium (SPSE)

the outside influence $\overline{N}$ is small and elasticity of substitution between ideas σ is large. Figure 2 illustrates the set of parameters σ and $\overline{N}$ in which the equilibrium characterized by Proposition 2 exists.

Note that $A(\overline{N})$ consists of two components, both of which are positively related to $\overline{N}$. The first component reflects the fact that with increasing outside influence, 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 outside influence 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 family size and increase the quality of own idea.

Another parameter of interest is the growth rate of the quality of ideas, which can be interpreted as the technological growth rate. From (7) and (9),

$$\frac{q_{t+1}}{q_t} = \left(\frac{E_{t+1}}{E_t}\right)^{\frac{1}{\sigma-1}} = \overline{x}(\overline{N})\left(1 + N\tau^{-(\sigma-1)}\right)^{\frac{1}{\sigma-1}} \tag{22}$$

where, to recall, $\overline{N} = N\tau^{-\sigma}$. As the learning cost τ falls, the rate of knowledge growth accelerates for two reasons. First, with better access to the knowledge of non-parents, agents are able to learn more given the same learning time.[3] Second, falling learning costs induce people to spend more time on learning $\overline{x}$, which also accelerates the rate of improvement of ideas.

[3]This result is similar to the famous "love of variety" effect in international trade literature.

4.3 Equilibrium stability

We next verify that the equilibrium derived in section 4.2 is locally stable. Two concepts of stability are introduced. *Intra*generational stability occurs when a representative individual responds to a small deviation of everyone else's learning choice x so that the average moves back to the original equilibrium. Formally,

$$\frac{dz}{dx}(\overline{x}, \overline{N}) < 1 \tag{23}$$

*Inter*generational stability occurs when after a small deviation in the learning environment of $j \in G_t$ (holding the learning environments of others in G_t unchanged), learning environments of j's successors converge to that of their contemporaries. Formally, the operator $\mathfrak{R}$ that maps the relative learning environment of j, $R(j) \equiv \frac{E(j)}{E_t}$, into the relative learning environment of his children, $R(C(j)) \equiv \frac{E(C(j))}{E_{t+1}}$, is the contraction mapping in the vicinity of unity:

$$\mathfrak{R}(1) = 1 \quad \text{and} \quad \left| \frac{d\mathfrak{R}}{dR}(1) \right| < 1 \tag{24}$$

The following Proposition establishes both types of stability.

Proposition 3 *The symmetric equilibrium characterized in Section 4.2 is intragenerationally and intergenerationally stable.*

The proof is contained in Appendix C.

4.4 The pure 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. Each agent $j \in G_t$, either artist or peasant, can learn from two sources. First, he can learn from own parent $C^{-1}(j)$, with the unitary learning cost as before. Second, he can learn from the mass N of the nearest artists; denote the set of these artists by $N(j)$. The cost of learning from artists in $N(j)$ is $\tau \geq 1$. Learning from other individuals is impossible. Note that when $\alpha = 0$, we are back to the isolated families scenario analyzed in section 4.1. When $\alpha = 1$, we are in the symmetric influence equilibrium analyzed in section 4.2.

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 can be shown to be an incentive compatible strategy.[4] 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 $E(j) = E_t$ which simplifies the analysis.

We now proceed to the analysis of learning and fertility decisions made by peasants. Since they can influence only own children, the second component of their utility (11) is zero. Since the only outside influence their children face is from artists who learn $x = L$ units of time, the utility of peasants can be rewritten as

$$U(z, \overline{N}) = \left(\frac{L-z}{\nu} \right) B(z, L, \overline{N}) \tag{25}$$

[4] Intuitively, the mass of individuals each artist can influence is inversely related to the share of artists in the population α. When α approaches zero, the slope of the second component of artist's utility (11) approaches infinity, which makes the artist maximize the learning time and have no family.

It is trivial to show that there exists a unique interior maximum of (25), hence the first derivative of utility must to equal to zero at the optimal point. Modifying (14), we obtain:

$$\nu U_z(z, \overline{N}) = B(z, L, \overline{N})F(z, L, \overline{N}) = 0 \tag{26}$$

Since B is positive for $z > 0$, we have that $F(z, L, \overline{N}) = 0$. The response of peasants to increase of outside influence $\overline{N}$ can be computed from the implicit function theorem:

$$\frac{dz(\overline{N})}{d\overline{N}} = -\frac{\frac{dF(z,L,\overline{N})}{d\overline{N}}}{\frac{dF(z,L,\overline{N})}{dz}}$$

From the definition of F (15), it is straightforward to verify that $\frac{dF(z,L,\overline{N})}{d\overline{N}} > 0$ and that $\frac{dF(z,L,\overline{N})}{dz} < 0$, and therefore peasants respond to increases in outside influence by acquiring more knowledge z and by reducing the family size $\frac{L-z}{\nu}$.

5 Outside influence vs. economic change: empirics

Identifying the effect of cultural influence on fertility is difficult because of the lack of credible measures of cultural influence. In this section, I refer to existing empirical and anecdotal evidence that can be interpreted as evidence of such an effect, and conduct a cross-country empirical study of my own that suggests the existence of such influence.

5.1 The effect of television on fertility

A stark example of the effect of outside influence is a recent finding by La Ferrara et al. [2008] who discover that the arrival of a TV channel broadcasting "soap operas" has a negative impact on fertility in Brazil. It is hard to think of ways in which "soap operas" could have improved the return to human capital, or social security, or child survival rate; an increased outside cultural influence is a better theory.

5.2 Religious isolation and fertility

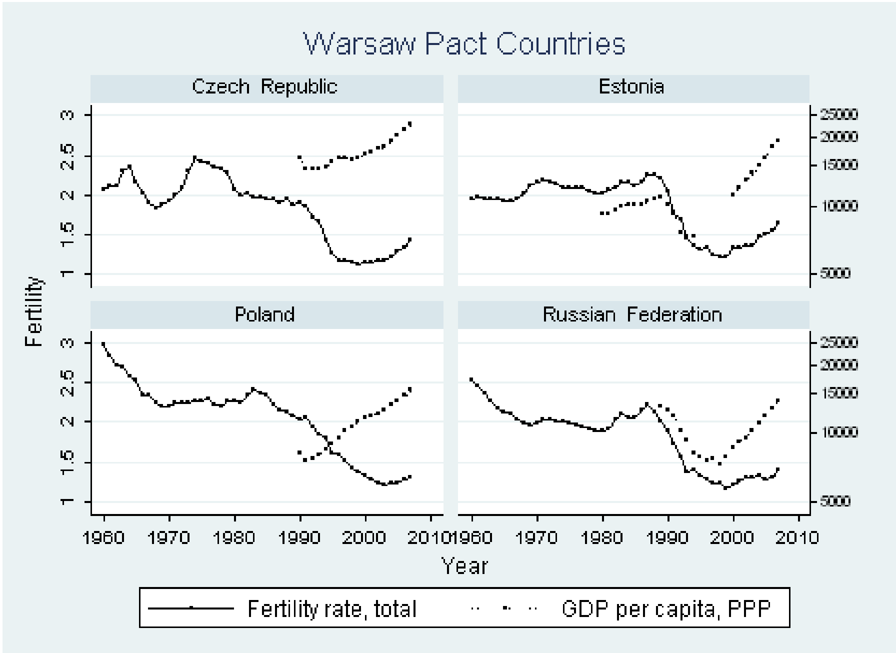
A good example of the opposite phenomenon – little contact with the outside world and preserved high fertility – is Amish community in the United States. They voluntarily abstain from modern communication devices such as phones and television, as well as from modern vehicles and travel opportunities. 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 [marital] fertility among Old Order Amish at 7.7 children per woman, which makes Amish community one of the fastest-growing communities in the OECD countries. As citizens of the United States, they have access to all economic opportunities that their non-religious compatriots do; the levels of their education, infant survival rates, and old age security presumably exceed those of people in the lower income decile of the US population, who are not known to have exceptionally high fertility rates. Isolation from the outside influence, rather than economic conditions, appears to be a better explanation of high Amish fertility.

5.3 The noxious influence of the West

A well-known stylized fact is that fertility in the Eastern European countries declined sharply in the late 1980s and early 1990s, following political and economic liberalization. For example, fertility (births per woman) in Russia fell from 2.22 in 1987 to 1.23 in 1997, almost a twofold decline in one decade. To my knowledge, there is no formal theory of the determinants of such a sharp decline; economic instability is usually blamed. All socioeconomic theories of demographic transition, however, associate macroeconomic instability with higher, not lower, fertility. Decreased return to human capital,[5] retirement system destroyed by hyperinflation, and increased infant mortality must have changed fertility in the opposite direction. Moreover, in a cross-section of countries, fertility decline does not seem to be closely related to macroeconomic turbulence. Fertility fell sharply in Czech Republic which has experienced only a mild and short recession; at the same time, there was no change of fertility trend in Tajikistan after a four-fold decline of per capita GDP. Serbia lived through a two-fold GDP decline, two civil wars and one war with NATO, which did not alter its fertility trend. An investigation of another region that lived through large macroeconomic turbulence in the 1980s and 1990s – Latin America – also does not reveal any relation between macroeconomic conditions and fertility. Argentina had a period of hyperinflation and a recession in late-1980s, which had no effect on the fertility trend.

The above mentioned evidence can be summarized as follows: fertility fell in countries of the former Warsaw pact, i.e. those behind the "iron curtain", except in countries with traditional "Oriental" culture like Azerbaijan or Uzbekistan. There was no sharp change in

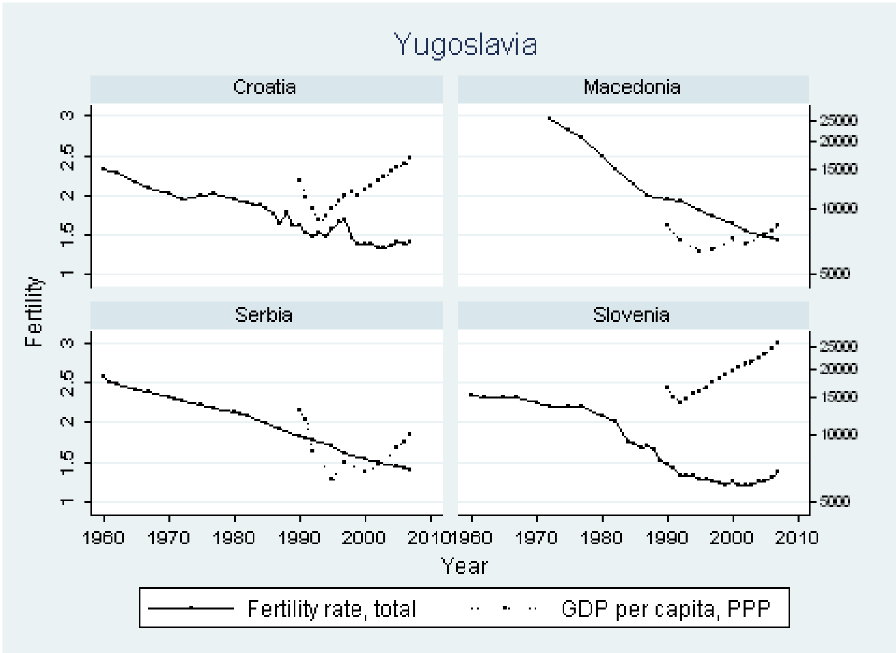
[5]E.g. the wages of university professors have notoriously fallen several times below the wages of taxi drivers, making the return to human capital effectively negative



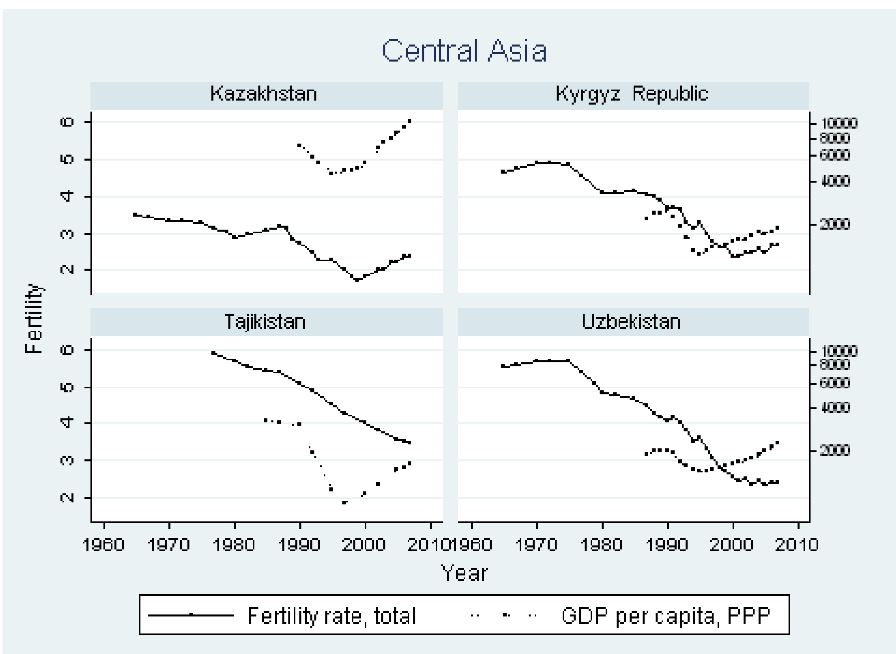
fertility in countries outside of the "iron curtain", e.g. former Yugoslav republics or Latin American countries, despite comparable macroeconomic turbulence.

These stylized facts can be explained by the "influential" theory of demographic transition. Prior to political reforms, members of the Warsaw pact were culturally isolated from the West. Soviet propaganda widely used the term "tletvornoye vliyanie Zapada" (the noxious influence of the West); a considerable effort was made to reduce the "noxious" influence to a minimum. In the late 1980s, all checks on the "noxious" influence have been removed; Eastern European countries got flooded by Western culture: music, movies, fashion, and lifestyles. A sharp increase in the outside cultural influence induced the young people in countries like Estonia, Russia and Moldova to acquire new types of "human capital" that would enable them to preserve their social status and influence in their communities; more time spent on acquisition of new human capital meant less time remaining for family.

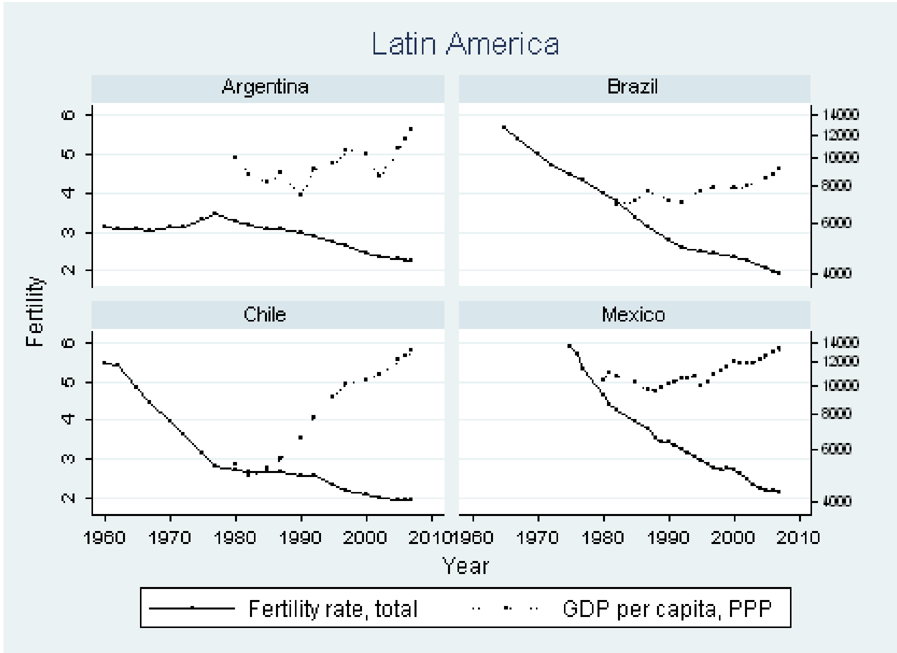
There was no change in fertility trends in former Yugoslavia and Latin America, because these territories were previously open to Western cultural influence, and large macroeconomic changes were not accompanied by large cultural change. There was no change in fertility trends in former Soviet countries with Oriental culture, because these countries were not susceptible to Western influence. To explain formally fertility non-response of Oriental cultures to a sharp rise in the Western influence, one could generalize the model developed in this paper in the following way. Two types of ideas can be introduced – Western and



Yugoslavia
Croatia
Macedonia
Serbia
Slovenia
Fertility
Year
Fertility rate, total
GDP per capita, PPP



Central Asia
Kazakhstan
Kyrgyz Republic
Tajikistan
Uzbekistan
Fertility
Year
Fertility rate, total
GDP per capita, PPP



Oriental; the production function of new ideas (2) would be characterized by high elasticity of substitution between ideas of different types. Then, individuals living in the Orient, surrounded by neighbors with predominantly Oriental ideas, would demand little of Western ideas; their observed behavior would change little following large changes in the access to Western ideas.

5.4 Fertility and happiness

The common property of all socioeconomic theories of demographic transition is that they associate reduced fertility with better economic opportunities or, in the language of undergraduate Microeconomics, with an expansion of the opportunity set of parents. A corollary of this assumption is that lower fertility must be accompanied by a higher level of utility. Increased return to human capital lowers fertility *and* increases utility. Increased old-age security lowers fertility *and* increases utility. Reduced child mortality lowers fertility *and* increases utility.

The "influential" theory of demographic transition developed in this paper highlights two forces that reduce fertility: increased opportunity to influence others, which increases utility, and increased opportunity of others to influence one's children, which lowers utility. It is natural to assume that in the era of mass-media, the second effect dominates for the vast majority of people, and hence the "influential" theory of demographic transition predicts a *positive* relationship between fertility and the level of utility: increased outside influence lowers fertility and *lowers* utility, despite the acceleration of human capital accumulation.

Empirically, the relationship between fertility and utility may be tested by using a measure of well-being or happiness as a proxy for utility. In the existing literature, Blanchflower and Oswald [2004] famously discover that the happiness trend in the US and Britain after the World War II (when fertility was declining) is negative, which is consistent with the theory of increasing outside influence following the spread of mass-media. Guriev and Zhuravskaya [2009] document that people in transition countries, especially people over 40 years old, are abnormally unhappy. This phenomenon, combined with abnormally low fertility in these countries discussed in section 5.3, can be explained by a sudden increase of the outside cultural influence that has decreased the influence of local community members on each other. The loss of social influence of older people was especially rapid, because, unlike the young, they had no opportunity to adjust their family size and absorb newly available influential ideas to retain own influence.

The relationship between well-being and utility can be tested directly whenever a measure of well-being is available. The underlying theory is as follows. Suppose the utility of a representative individual is $u(f - \epsilon, X, Z) + \nu$, where f is fertility chosen by the decision maker, X is $1 \times n_X$ vector of observed exogenous utility determinants, Z is $1 \times n_Z$ vector of unobserved exogenous utility determinants (including outside influence), and ϵ, ν are shocks that are observed by the individual but not by the econometrician, and that are orthogonal to X and Z.[6] The utility is smooth and concave, and therefore results in the unique optimal fertility level $f(x, z) - \epsilon$. By the envelope theorem, therefore, full derivatives of utility with respect to control variables (that is, the marginal effects of controls on utility) are equal to the corresponding partial derivatives; the derivative of (optimal) utility with respect to (optimal) fertility is zero. The utility at the optimal point can therefore be linearized as follows:[7]

$$u(f(X, Z) - \epsilon, X, Z) = X\alpha_X + Z\alpha_Z + \nu \qquad (27)$$

while the optimal fertility can be linearized as

$$f(X, Z) = X\beta_X + Z\beta_Z + \epsilon \qquad (28)$$

[6] We assume unconstrained optimization, hence u can be viewed as the Lagrangian of the corresponding constrained optimization problem.

[7] All variables are demeaned for tractability of exposition.

The unobserved controls Z can be inferred from (27) as

$$Z = (u - X\alpha_X - \nu)\alpha_Z^{-1} \tag{29}$$

where α_Z^{-1} is $1 \times n_X$ vector of inverse elements of α_Z. Then, fertility can be rewritten as follows:

$$f = X(\beta_X - \alpha_X\alpha_Z^{-1}\beta_Z) + u\alpha_Z^{-1}\beta_Z - \nu\alpha_Z^{-1}\beta_Z + \epsilon \tag{30}$$

By regressing f on X and u, we obtain $\alpha_Z^{-1}\beta_Z$ as a coefficient for u.[8] If it is positive and significant, there must exist a element of Z which is changing fertility and utility in the same direction, which we attribute to the effect of outside cultural influence due to lack of alternative theories making such a prediction.

I test this prediction by using data from the World Values Survey, which is a globally representative sociological survey conducted in almost 100 countries of the world, with five waves spanning the years of 1980-2005. Among other questions, the survey asks its respondents to rate the overall satisfaction with their lives on the scale from 1 to 10, which I use as a proxy for utility. Although the data on the actual number of children is available, I avoid using it because actual fertility outcomes may depend on unobserved health shocks which make an individual simultaneously less fertile and less satisfied with life, which would make the regression results spurious. Instead, as a measure of fertility I choose the *ideal* number of children reported by the respondents. The ideal number of children is, arguably, less related to idiosyncratic shocks such as health, unemployment, and widowhood, and is more likely to be an outcome of the overall cultural and socioeconomic environment. There are 166 thousand observations in about 65 countries on all continents in which both life satisfaction and desired fertility are observed. Table 1 in Appendix reports the results.

At the global level, the effect of utility, as a proxy for unobserved determinants of fertility, is positive and significant as long as country dummies are included into the regression. This effect is positive on all continents except Africa, and is significant in America, Asia, and Europe. In the Pacific (Australia and New Zealand), it is of the same order of magnitude as on the preceding three continents, but a smaller number of observations do not allow to judge about its significance. Therefore, we can conclude, the variation of fertility that was not explained by cross-country differences is more plausibly explained by variation in outside cultural influence than by variation in socioeconomic conditions.

The fact that without country fixed effects the relationship between fertility and well-

[8]The estimate of this coefficient is closer to zero than the true value due to error-in-variable ν.

being is negative can be explained as follows. While low outside influence, all else equal, results in larger and happier households, high fertility in such territories will reduce the nationwide amount of resources per capita, which lowers utility. With this theory, we are more likely to observe the positive relationship between well-being and fertility within countries than across countries.

6 Conclusion

This paper develops an "influential" theory of demographic transition, according to which the decline of fertility over time occurs due to increased outside cultural influence rather than due to increased economic opportunities of parents. The objective of parents is to maximize own cultural influence on own children and on other people. With increased outside influence, children's demand for parents' ideas becomes more sensitive to their quality, which induces parents to spend more time on education and reduce the amount of time spent on family. It is difficult to imagine a father of seven children attend a rock concert or play in a rock band; rock music, almost by definition, is a high-influence entity. Access of people to this kind of activities should have an effect on fertility choices.

The model developed in this paper, for simplicity, assumes that ideas are unidimensional; in fact, there may exist a vast array of different types of ideas. Different societies may be susceptible to different types of ideas, and it may be hard to tell in advance which ideas will affect a particular society. La Ferrara et al. [2008] discover that fertility in Brazil declines after the arrival of a particular television channel broadcasting particular types of shows; the same channel is likely to have little effect on African households. In this light, the Internet seems to be the best vector of influential ideas, as it allows users to search and find themselves the ideas that influence them most.

The model developed in this paper predicts that high outside influence, besides reducing fertility, makes people worse off. However, reduced fertility may have a positive externality on other families by increasing the amount of resources per capita. Therefore, the aggregate effect of increased outside influence on well-being may be positive. The empirical part of the paper confirms that within countries, the relationship between fertility and a measure of utility is positive, while across countries it is negative.

References

Gary S. Becker. An economic analysis of fertility. In *Demographic and Economic Change in Developed Countries*, pages 225 – 256. UMI, 1960.

Gary S. Becker and H. Gregg Lewis. On the interaction between the quantity and quality of children. *The Journal of Political Economy*, 81(2):S279–S288, 1973.

Gary S. Becker, Kevin M. Murphy, and Robert Tamura. Human capital, fertility, and economic growth. *The Journal of Political Economy*, 98(5):S12–S37, 1990.

Alberto Bisin and Thierry Verdier. On the cultural transmission of preferences for social status. *Journal of Public Economics*, 70(1):75 – 97, 1998.

Alberto Bisin and Thierry Verdier. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, 97(2):298 – 319, 2001.

Alberto Bisin, Giorgio Topa, and Thierry Verdier. Religious intermarriage and socialization in the United States. *Journal of Political Economy*, 112(3):615–664, 2004.

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

David G. Blanchflower and Andrew J. Oswald. Well-being over time in Britain and the USA. *Journal of Public Economics*, 88(7-8):1359 – 1386, 2004.

Robert Boyd and Peter J. Richerson. *The Origin and Evolution of Cultures*. Oxford University Press, 2005.

John C. Caldwell. Toward a restatement of demographic transition theory. *Population and Development Review*, 2(3/4):321–366, 1976.

Harold L. Cole, George J. Mailath, and Andrew Postlewaite. Social norms, savings behavior, and growth. *Journal of Political Economy*, 100(6):1092, 1992.

Richard Dawkins. *The Selfish Gene*. Oxford University Press, 1976.

C. Simon Fan and Oded Stark. Looking at the population problem through the prism of heterogeneity: Welfare and policy analyses. *International Economic Review*, 49:799–835(37), August 2008.

Oded Galor and Omer Moav. Natural selection and the origin of economic growth. *Quarterly Journal of Economics*, 117(4):1133 – 1191, 2002.

Oded Galor and David N. Weil. The gender gap, fertility, and growth. *The American Economic Review*, 86(3):374–387, 1996.

Oded Galor and David N. Weil. Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. *The American Economic Review*, 90(4):806–828, 2000.

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

Sergei Guriev and Ekaterina Zhuravskaya. (Un)Happiness in Transition. *The Journal of Economic Perspectives*, 23:143–168(26), 2009.

W. D. Hamilton. The genetical evolution of social behaviour, I. *Journal of Theoretical Biology*, 7(1):1 – 16, 1964.

Joseph Henrich 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*, 22(3):165 – 196, 2001.

Sebnem Kalemli-Ozcan. A stochastic model of mortality, fertility, and human capital investment. *Journal of Development Economics*, 70:103–118(16), 2003.

Hans-Peter Kohler, Francesco C. Billari, and Jose Antonio Ortega. The emergence of lowest-low fertility in Europe during the 1990s. *Population and Development Review*, 28(4): 641–680, 2002.

Eliana La Ferrara, Alberto Chong, and Suzanne Duryea. Soap operas and fertility: Evidence from Brazil. RES Working Papers 4573, Inter-American Development Bank, Research Department, June 2008.

Kaivan Munshi and Jacques Myaux. Social norms and the fertility transition. *Journal of Development Economics*, 80(1):1 – 38, 2006.

Lesley Newson, Tom Postmes, S.E.G. Lea, Paul Webley, Peter J. Richerson, and Richard McElreath. Influences on communication about reproduction: the cultural evolution of low fertility. *Evolution and Human Behavior*, 28(3):199 – 210, 2007.

Vasily Peskov. *Lost in the Taiga: One Russian Family's Fifty-Year Struggle for Survival and Religious Freedom in the Siberian Wilderness*. Doubleday, 1994.

Arthur J. Robson. Status, the distribution of wealth, private and social attitudes to risk. *Econometrica*, 60(4):837–857, 1992.

Arthur J. Robson. The biological basis of economic behavior. *Journal of Economic Literature*, 39(1):11–33, 2001.

Robert Tamura. Human capital and economic development. *Journal of Development Economics*, 79(1):26 – 72, 2006.

World Values Survey 1981-2008. World Values Survey Association, official aggregate v.20090901 edition, 2009.

A Derivation of demand for ideas

Maximization of (2) subject to (3) over all learning intensities $y(i,j)$ is equivalent to the maximization of the following Lagrangian:

$$\max_{y(i,j),\forall i\in\{N(j)\cup C^{-1}(j)\}} \left\{ q(j) - \lambda(j)\left[x(j) - y(C^{-1}(j),j) - \int_{i\in N(j)} y(i,j)\tau di \right] \right\} \tag{31}$$

where $q(j)$ is defined in (2), and $\lambda(j)$ is the Lagrange multiplier. Maximization with respect to an arbitrary $y(i,j)$ yields

$$q(j)^{\frac{1}{\sigma-1}} y(i,j)^{-\frac{1}{\sigma}} q(i)^{\frac{\sigma-1}{\sigma}} - \lambda(j)\tau^{I(j\notin C(i))} = 0 \tag{32}$$

Solving for $y(i,j)$ yields:

$$y(i,j) = K(j)q(i)^{\sigma-1}\tau^{-\sigma I(j\notin C(i))} \tag{33}$$

where $K(j)$ is some positive person-specific constant. We can now substitute (33) into (3) to obtain

$$K(j)q(C^{-1}(j))^{\sigma-1} + \int_{i\in N(j)} K(j)\left(\frac{q(i)}{\tau}\right)^{\sigma-1} di = x(j) \tag{34}$$

By using the definition of the learning environment (5), we can simplify the above formula to $K(j)E(j) = x(j)$, or $K(j) = \frac{x(j)}{E(j)}$. By substituting the latter expression into (33), we end up with the optimal learning intensities presented in (6).

B Proof of Lemma 1

Throughout the proof, the function arguments x and $\overline{N}$ are held fixed and are suppressed in the presentation.

Define a function $W(z)$ such that (cf.(14))

$$W_z(z) = \nu z^{-(\sigma-2)}U_z(z) = z^{-(\sigma-2)}B(z)F(z) + K \tag{35}$$

where K is a positive constant. By definition, W_z has the same sign as U_z for all $z > 0$, and therefore a point $z > 0$ is a local maximum (minimum) of $W(z)$ if and only if it is the point of a local maximum (minimum) of $U(z)$. We focus on the properties of $W(z)$ in the following analysis.

By definition of B and F (cf.(12),(15)), $W(z)$ is strictly increasing on $(0, z^*]$, where z^* is defined in (16). Hence, we focus on studying W on $z \in (z^*, L]$. We start by investigating the sign of the second derivative of W. In equations that follow, all function arguments are suppressed for brevity.

$$
\begin{aligned}
W_{zz} &= -(\sigma-2)z^{-(\sigma-1)}BF + (\sigma-1)z^{-(\sigma-1)}B(1-B)F \\
&- z^{-(\sigma-1)}B(\sigma-1)\frac{L}{z}(1-B) - z^{-(\sigma-1)}B(\sigma-1)(\frac{L}{z}-1)(\sigma-1)B(1-B)
\end{aligned}
$$

Division of W_{zz} by $Bz^{-(\sigma-1)}$ will change neither its sign nor its zeros. We obtain

$$
\begin{aligned}
B^{-1}z^{\sigma-1}W_{zz} &= -(\sigma-2)F + (\sigma-1)(1-B)F \\
&- (\sigma-1)\frac{L}{z}(1-B) - (\sigma-1)\left(\frac{L}{z}-1\right)(\sigma-1)B(1-B)
\end{aligned} \tag{36}
$$

Next, I add $(\sigma-1)B + (\sigma-1)(1-B) - (\sigma-1)$, which equals zero, to the right-hand side of (36):

$$
\begin{aligned}
B^{-1}z^{\sigma-1}W_{zz} &= -(\sigma-2)F + (\sigma-1)(1-B)F - (\sigma-1)\left(\frac{L}{z}-1\right)(1-B) \\
&- (\sigma-1)B\left((\sigma-1)\left(\frac{L}{z}-1\right)(1-B)-1\right) - (\sigma-1) \\
&= -(\sigma-2)F + (\sigma-1)(1-B)F - (F+1) - (\sigma-1)BF - (\sigma-1) \\
&= -2(\sigma-1)BF - \sigma
\end{aligned} \tag{37}
$$

By definition of z^*, $F(z)$ is strictly negative on $(z^*, L]$ and is decreasing from 0 to -1, therefore (37) is strictly increasing from $-\sigma$ to $2(\sigma-1)B(L,x,R)-\sigma$ on $(z^*, L]$. Whenever (17) holds, W_{zz} is strictly negative for any value of $z \in (z^*, L)$; negative second derivative implies concavity of W and uniqueness of argmaximum, which proves the third statement of Lemma 1. When (17) does not hold, there exists $z^{**} \in (z^*, L)$ such that $W_{zz}(z) < 0$ for $z \in (z^*, z^{**})$ and $W_{zz}(z) > 0$ for $z \in (z^{**}, L)$. Then, there may be at most one local maximum of W on $z \in (z^*, z^{**})$; another possible maximum is at $z = L$. This proves the first and second statements of Lemma 1. ■

C Proof of Proposition 3

C.1 Intragenerational stability

Given that the equilibrium derived in Section 4.2 is interior (i.e. $z(\overline{x}, \overline{N}) \in (0, L)$), the best response function in the vicinity of the equilibrium is found from the condition $U_z(z, x, \overline{N}) = 0$. From the implicit function theorem,

$$\frac{dz(x, \overline{N})}{dx} = \frac{U_{zx}(\overline{x}, \overline{x}, \overline{N})}{-U_{zz}(\overline{x}, \overline{x}, \overline{N})}$$

At the point of utility maximum, the denominator is positive (cf.(21)), and therefore the sign of $\frac{dz(\overline{x}, \overline{N})}{dx}$ is equal to the sign of $U_{zx}(\overline{x}, \overline{x}, \overline{N})$:

$$U_{zx}(\overline{x}, \overline{x}, \overline{N}) = -\frac{\sigma - 1}{\overline{x}} \left[\frac{\overline{N}^2}{A(\overline{N})(1 + \overline{N})^3} + \frac{\overline{N}}{1 + \overline{N}} \right] < 0 \tag{38}$$

Therefore, when others change their learning choice, the decision maker responds by changing his decision in the opposite direction, moving the system towards the initial equilibrium. ■

C.2 Intergenerational stability

Throughout most of the paper, it is assumed that learning environments are the same for all members of a given generation. This section emphasizes that the learning environment of an individual may differ from that of his contemporaries: $R(j) = \frac{E(j)}{E_t} \neq 1$, which will alter the choices of that individual. To take account of that, I redefine utility as a function of the learning choice and the relative learning environment (cf.(10)):

$$U(z, R) = \frac{L - z}{\nu} \frac{R z^{\sigma - 1}}{R z^{\sigma - 1} + \overline{x}^{\sigma - 1} \overline{N}} + \frac{L - \overline{x}}{\nu} \frac{R z^{\sigma - 1} \overline{N}}{\overline{x}^{\sigma - 1} + \overline{x}^{\sigma - 1} \overline{\overline{N}}} \tag{39}$$

The learning choice of others $\overline{x}$ is assumed to be at the equilibrium characterized in Proposition 2. We denote by $z(R)$ the optimal learning choice of the decision maker: $z(R) = \arg\max_z U(z, R)$. From the definition of $\overline{x}$ (cf.(18)), it follows that $z(1) = \overline{x}$.

With new notations, the relative learning environment of children of j is (cf.(9))

$$\begin{aligned}\mathfrak{R}(R(j)) \quad &\equiv \quad R(C(j)) = \frac{E(C(j))}{E_{t+1}} = \frac{E(j)z(R(j))^{\sigma-1} + NE_t\left(\frac{\overline{x}}{\tau}\right)^{\sigma-1}}{E_t\overline{x}^{\sigma-1} + NE_t\left(\frac{\overline{x}}{\tau}\right)^{\sigma-1}} \\ &= \quad \frac{R(j)\left(\frac{z(R(j))}{\overline{x}}\right)^{\sigma-1}}{1+N\tau^{-(\sigma-1)}} + \frac{N\tau^{-(\sigma-1)}}{1+N\tau^{-(\sigma-1)}} \end{aligned} \tag{40}$$

From (40), it follows directly that $\mathfrak{R}(1) = 1$. By differentiating (40) with respect to R, we obtain

$$\frac{d\mathfrak{R}(1)}{dR} = \frac{1+(\sigma-1)\frac{z_R(1)}{\overline{x}}}{1+N\tau^{-(\sigma-1)}} = \frac{1+\frac{U_{zR}(\overline{x},1)}{-\frac{\overline{x}}{\sigma-1}U_{zz}(\overline{x},1)}}{1+\overline{N}\tau} \tag{41}$$

It can be shown that $U_{zR}(\overline{x},1) = \frac{\overline{N}^2}{(1+\overline{N})^3 A(\overline{N})} > 0$ and therefore $\frac{d\mathfrak{R}(1)}{dR} > 0$. To prove intergenerational stability, it is sufficient to prove that $\frac{d\mathfrak{R}(1)}{dR} < 1$. Indeed, we have that

$$\frac{U_{zR}(\overline{x},1)}{-\frac{\overline{x}}{\sigma-1}U_{zz}(\overline{x},1)} = \frac{\overline{N}}{(1+\overline{N})\left(N+\frac{2+N}{\sigma-1}\right)} \tag{42}$$

Intergenerational stability holds if (42) is strictly less than $\overline{N}\tau$, or

$$\frac{1}{\tau} - (1+\overline{N})\left(N+\frac{2+N}{\sigma-1}\right) < 0 \tag{43}$$

Although (43) does not hold for all values of τ, $\overline{N}$ and σ, it can be shown that it does hold for any $\tau \geq 1$ if (20) holds. In other words, if the symmetric pure strategy equilibrium exists, it is intergenerationally stable. ■

Table 1: The Observed Effect of Life Satisfaction on Desired Family Size

Ordered Probit Regression

Dependent variable: ideal number of children

	World				Africa	America	Asia	Europe	Pacific
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Life satisfaction	-0.0027**	0.0055**	0.0109**	0.0110**	-0.0195**	0.0143**	0.0076**	0.0205**	0.0126
	(0.0010)	(0.0011)	(0.0013)	(0.0013)	(0.0026)	(0.0027)	(0.0021)	(0.0024)	(0.0084)
Age			0.0100**	0.0100**	0.0127**	0.0118**	0.0151**	0.0108**	0.0063**
			(0.0002)	(0.0002)	(0.0005)	(0.0004)	(0.0004)	(0.0003)	(0.0009)
Education dummies			Yes	Yes					
Income decile dummies			Yes	Yes					
Country dummies		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey wave dummies			Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-wave interaction dummies				Yes					
Observations	165983	165983	126558	126558	21879	33527	47830	45155	4787

Standard errors in parentheses

* significant at 5%; ** significant at 1%