# Looking at the "Population Problem" through the Prism of Heterogeneity: Welfare and Policy Analyses

*By* C. SIMON FAN and ODED STARK

# LOOKING AT THE "POPULATION PROBLEM" THROUGH THE PRISM OF HETEROGENEITY: WELFARE AND POLICY ANALYSES*

By   C. Simon Fan and Oded Stark[1]

*Lingnan University, Hong Kong; University of Klagenfurt, Austria, University of Bonn, Germany, University of Vienna, Austria, University of Warsaw, Poland, and Warsaw School of Economics, Poland*

Upon introducing heterogeneity and dynamics into a model of the demand for children, a problem of optimal population is defined and analyzed. It is shown that from the perspective of social welfare, better-educated individuals produce too few children while less-educated individuals produce too many children and all individuals invest too little in the education of their children. The impact of several policy tools geared at addressing the "population problem" is investigated, in particular how child allowances and other tax-subsidy policies can be harnessed to enhance welfare, and how and why early childhood education programs can mitigate the "population problem."

## 1. INTRODUCTION

Much of the economics literature has identified "the population problem" as a major cause of the "poverty trap" in economic development.[2] In particular, de la Croix and Doepke (2003) demonstrate that due to the negative correlation between income and fertility, and the positive correlation between income and investment in children's human capital, an increase in inequality lowers the average level of education and depresses the growth rate.[3] Utilizing a framework of differential fertility, de la Croix and Doepke (2004) assess the merits of different public policies. An intuitive and straightforward implication of the illuminating writings of de la Croix and Doepke seems to be that a low-income household produces too many children and invests too little in each child's human capital, although de la Croix and Doepke do not explicitly analyze social welfare.

[2] Notable contributions of this literature include Dahan and Tsiddon (1998), Galor and Weil (2000), Galor and Moav (2002, 2004, 2006), Kremer and Chen (2002), de la Croix and Doepke (2003, 2004), Chakraborty and Das (2005), Moav (2005), and Soares (2005). Dasgupta (1995), Ehrlich and Lui (1997), and de la Croix and Michel (2002) provide useful surveys of this literature.

[3] De la Croix and Doepke (2003) show that this fertility-differential effect accounts for most of the empirical relationship between inequality and growth.

The current inquiry seeks to complement de la Croix and Doepke by conducting a welfare and policy analysis of the population problem under the assumption of heterogeneity. Our inquiry is also related to the received literature on "the optimal population," which is mostly based on the "representative individual" framework.[4] We extend this literature by considering differential fertility with heterogeneous agents in the analysis of social welfare. The purpose of this extension is to improve the understanding of the population problem in an economy in which the distribution of income and of human capital is uneven. Our analytical framework is in line with a recent contribution by Mookherjee and Napel (2007), who develop a dynamic model of heterogeneous agents that allows for both downward and upward intergenerational mobility. However, whereas Mookherjee and Napel assume that fertility is fixed, we analyze the interactions between the quantity and the quality of children, in line with Becker and Lewis (1973), and more recently in line with de la Croix and Doepke.

Our analysis focuses on the "steady state" properties of the population dynamics. As in Mookherjee and Napel (2007), we define a steady state as a state wherein the proportion of skilled workers in the economy is not changing. Meanwhile, in our model, the population growth rate also remains constant in a steady state. The model examines the existence, uniqueness, and stability of the steady state. Upon incorporating a Becker-Lewis type model into our framework of population dynamics and then drawing on the steady-state characterization, we analyze and characterize a "population problem" with heterogeneous agents. We show that based on a social welfare function commonly used in the received literature, better-educated individuals produce too few children while less-educated individuals produce too many children *and* all individuals invest too little in the education of their children.

Drawing on this infrastructure, we examine policies that could correct or mitigate the "population problem." We study a tax and subsidy policy aimed at increasing the number of children of skilled individuals and at raising social welfare. We find (analytically) that such a policy is riddled with implementation difficulties: Either the number of children of the skilled individuals will increase yet their investment in the skill acquisition of each child will decrease, or the inverse. Moreover, the welfare of the skilled individuals decreases *and* the policy's impact on social welfare is ambiguous. We then examine the efficacy of a program aimed at improving the skill-formation environment in which the skills of the children of the unskilled parents are acquired. We find that, under a fairly mild condition, the program can mitigate the "population problem" and raise social welfare. Thus, our foray into the policy arena results in a somewhat surprising verdict: A policy that is expected to work can well result in outcomes that are in variance with the intended outcomes and a policy that is not aimed at addressing the "population problem" at all has the unintended consequence of effectively mitigating

[4] For example, see Samuelson (1975, 1976), Gigliotti (1983), Nerlove et al. (1984, 1985, 1986a, 1986b, 1988), Razin and Sadka (1995), Boucekkine et al. (2002), and Razin and Yuen (2004).

the problem.[5] Moreover, when we extend the model by considering the positive externality effect of human capital, we identify a scheme that can improve the welfare of the unskilled individuals and the net income of the skilled individuals in every generation. The scheme evolves around the government borrowing to pay the unskilled individuals of a given generation to reduce their fertility, while being paid back by the skilled individuals of a subsequent generation, with this procedure revolving generation after generation.

In the reminder of this paper, Section 2 presents a static, Becker-Lewis type model of fertility. Section 3 analyzes population dynamics with heterogeneous individuals in every generation. Section 4 addresses the "population problem" in the new framework developed in this paper. Sections 5–7 investigate several policy implications. Section 8 offers an inquiry into the validity of our main analytical results under alternative social welfare functions. Section 9 presents our conclusion. Proof material is relegated to the Appendix.

## 2. A SIMPLE MODEL OF FERTILITY

This section is a variant of a Becker-Lewis type model. Consider an economy that extends over an infinite discrete time. In every period, the economy is populated by a large number of individuals. An individual lives for two periods: childhood and adulthood (parenthood). The individual makes decisions only in the second period of his life. During this period, the individual bears children and chooses how much education to provide to his children. When the children mature, and having had their skills formed, they in turn choose how many children to have and how much education to give to their children. There are two types of individuals in the economy: skilled individuals—individuals who, for example, receive higher education—and unskilled individuals.

An adult obtains utility from three sources: consumption of material goods, $c$; the number of children, $n$; and an index for the "quality" of children. We measure the "quality" of children by the probability that a child will be skilled. Let $m$ be the probability that a child will be skilled. As noted by Razin and Sadka (1995), the Becker-Lewis model with general functional forms does not necessarily imply that richer people have fewer children. We therefore specify a certain functional form of the utility function. Let an adult's utility function be

$$(1) \qquad U(n, m, c) \equiv \ln n + \alpha \ln m + \beta \ln c,$$

where $\alpha$ and $\beta$ are positive coefficients.

A parent is endowed with one unit of time, which is devoted to (i) bearing and raising children; (ii) educating children; and (iii) working. In line with the existing literature, we assume that children (both quantity-wise and quality-wise)

are time intensive for their parents. Indeed, many empirical studies show that parents' time accounts for the bulk of the cost of producing and rearing children (see, for example, the survey by Becker, 1991). For the sake of simplicity and in line with Galor and Weil (2000), Galor and Moav (2002), and de la Croix and Doepke (2003, 2004), we assume that the only parental controlled input into the production of child quantity and quality is parents' time. Let $\xi$ and $k$ denote the time cost of bearing and raising a child and of educating a child, respectively. (We assume that parents spend the same amount of time educating each one of their children.) We assume that $\xi$ is a fixed parameter, while $k$ is a choice variable. Let a parent's income be $w$. Then,

$$(2) \qquad\qquad c = (1 - \xi n - kn)w$$

and the utility function (1) can be rewritten as

$$(3) \qquad\qquad \ln(n) + \alpha \ln(m) + \beta \ln[(1 - \xi n - kn)w].$$

Further, we define

$$m = e(k, \delta)$$

$$(4)$$
$$e_1 \equiv \frac{\partial e}{\partial k} > 0; \quad e_{11} \equiv \frac{\partial^2 e}{\partial k^2} < 0; \quad e_2 \equiv \frac{\partial e}{\partial \delta} > 0; \quad e_{12} \equiv \frac{\partial^2 e}{\partial k \partial \delta} > 0,$$

where $\delta$ is a coefficient that can be interpreted as an indicator of the home environment that helps shape the formation of human capital. It should be noted that $\delta$ is greater for skilled parents than for unskilled parents since, given the amount of time spent on educating a child, the children of skilled individuals who are exposed to a different value system, cultural values, and academic orientation are more likely to be skilled themselves.[6] Thus, the probability that a child will end up being skilled depends on parental investment and on parental type. To capture the idea that there are diminishing returns to the educational effort, we assume that $e_{11} \equiv \frac{\partial^2 e}{\partial k^2} < 0$. We also assume that the marginal benefit of education, $\frac{\partial e}{\partial k}$, increases with parental human capital ($\delta$), that is, $e_{12} \equiv \frac{\partial^2 e}{\partial k \partial \delta} > 0$, which means

---

[6] The importance of family background, and particularly of parental academic achievements and human capital, for an individual's educational attainment has been consistently confirmed in the empirical literature. For example, several educators (van der Eyken, 1977; Sahota, 1978) estimate that about 50 percent of cognitive development of children occurs by the age of 3 or 4. Bloom (1981) estimates that 80 percent of an individual's potential intelligence is developed by age 8. Clearly, as suggested by Osberg (1984), almost all this early conditioning environment is created by a child's parents, and it is in these early years that the basic personality traits of a child, such as self-confidence and achievement motivation, which significantly affect the child's future academic performance, are largely shaped. Some evidence (for example, Vernon, 1979) suggests that the home environment also affects children's efficiency in learning at school. Hanushek (1996) and de la Croix and Michel (2002) provide useful surveys of this and related literature.

that the parental time input and the parental trait are more complementary in the production of the quality of children.[7]

We assume that the optimal solutions are interior. The first order conditions for utility maximization are thus

$$(5) \qquad n : \frac{1}{n} = \frac{\beta(\xi + k)}{1 - \xi n - kn}$$

$$(6) \qquad k : \frac{\alpha e_1}{e} = \frac{\beta n}{1 - \xi n - kn}.$$

Now, we make the following assumption:

$$(7) \qquad \alpha(\xi + k)e_{12} - e_2 > 0.$$

Then, we have the following proposition.

PROPOSITION 1. *If (7) is satisfied, then*

$$\frac{dn}{d\delta} < 0; \frac{dk}{d\delta} > 0; \frac{dm}{d\delta} > 0.$$

PROOF.   See the Appendix.

For the proposition to hold, (7) is critical. Clearly, *ceteris paribus*, (7) is more likely to be satisfied the larger is $e_{12}$. The intuition is that a larger $e_{12}$ implies that the parental time input and the parental trait are more complementary in the production of the quality of children. Hence, we are more likely to have $\frac{dk}{d\delta} > 0$ when $e_{12}$ increases. We can also see that (7) will not be satisfied if $e_{12} \leq 0$. The intuition here is that if $e_{12} \leq 0$, then an increase in $k$ will reduce the marginal impact of the parental trait. Therefore, if $e_{12} \leq 0$, an increase in $\delta$, which increases $e(k, \delta)$, will lower $k$ and hence induce an increase in fertility at the optimum.

We now briefly turn our attention to the determination of the wage rates.[8] We assume that individuals operate in a small open economy in a one-good world. The good can be produced by two constant returns to scale production technologies. One production technology is "modern" wherein human capital (or skilled labor) is used intensively; the other production technology is "traditional" wherein physical (or unskilled) labor is used. To simplify, we assume that the only input in the modern (skilled) sector is efficiency labor (or "mental" labor), while the only input in the traditional (unskilled) sector is physical labor.

---

[7] Note that parental time input of educating children is a parent's opportunity cost of working, which, roughly speaking, implies that parental time input can be regarded as a parent's financial cost of educating children. As it is commonly assumed that parents' educational expenditures and parental human capital are complementary in children's education (Benabou, 1996; Galor and Tsiddon, 1997), this assumption is in line with the received literature.

[8] Note that the formulation here is similar to that in Galor and Zeira (1993) and Hazan and Berdugo (2002).

The production technologies in the traditional sector and in the modern sector are as follows. The production function of the traditional (unskilled) sector is

$$(8) \qquad Y_t^u = w_u L_t^u,$$

where $Y_t^u$ and $L_t^u$ are the output of this sector and the unskilled labor force of the entire economy at time $t$, respectively, and $w_u$ is the marginal productivity of physical labor. The market for physical labor is perfectly competitive. Thus, the wage rate of an unskilled worker is $w_u$.

The modern sector's production function is

$$Y = F(K, L) \equiv Lf(\vartheta), \qquad \vartheta \equiv \frac{K}{L},$$

where $Y$, $K$, and $L$ are total output, the quantity of physical capital, and the quantity of skilled labor, respectively. Since the economy is perfectively competitive, the interest rate of physical capital, $r$, and the wage rate of skilled labor, $w$, are determined as follows:

$$(9) \qquad \begin{aligned} r &= f'(\vartheta) \\ w &= f(\vartheta) - kf'(\vartheta). \end{aligned}$$

Suppose that the world interest rate is constant at $\bar{r}$. Assuming that the small economy permits unrestricted international lending and borrowing, its interest rate must also be equal to $\bar{r}$. Therefore, the ratio between capital and skilled labor in this economy is constant at a level $f'^{-1}(\bar{r}) \equiv \bar{\vartheta}$. Thus, the wage rate of the skilled workers, $w_s$, is constant at the level of

$$f(\bar{\vartheta}) - \bar{k}f'(\bar{\vartheta}).$$

In line with the empirical literature, we assume that income and skill are positively correlated, that is,

$$w_s > w_u$$

or that

$$\frac{dw}{d\delta} > 0.$$

Since there are only two types of individuals, we define

$$\delta = \begin{cases} 1 & \text{if the individual is skilled} \\ 0 & \text{otherwise.} \end{cases}$$

Then, from Proposition 1, we have the following corollary.

COROLLARY 1.  *Fertility is negatively correlated with income (with parental skill level) and the educational expenditure on each child increases with income (with parental skill level) if condition (7) is satisfied.*

### 3.   POPULATION DYNAMICS

To ease the exposition that follows we introduce the following notations:

$\lambda_t$: the proportion of skilled individuals in the population at time $t$;
$p$: the probability that a child whose parent is skilled will be skilled;
$q$: the probability that a child whose parent is unskilled will be skilled;
$a$: the fertility of a skilled individual,
$b$: the fertility of an unskilled individual,
$N_t$: the total size of cohort at time $t$.

From Proposition 1 we know that

$$p > q,$$

and that

$$a < b.$$

At time $t + 1$, the number of skilled individuals whose parents are skilled is

$$pa\lambda_t N_t.$$

At time $t + 1$, the number of skilled individuals whose parents are unskilled is

$$qb(1 - \lambda_t)N_t.$$

Therefore, the number of the skilled at time $t + 1$ is

$$pa\lambda_t N_t + qb(1 - \lambda_t)N_t.$$

Recalling our assumption that an adult lives (that is, makes economic decisions) for just one period, the total (adult) population at time $t + 1$ is

$$a\lambda_t N_t + b(1 - \lambda_t)N_t.$$

Thus, the proportion of the skilled at time $t + 1$ is

$$(10) \qquad \begin{aligned} \lambda_{t+1} &= \frac{pa\lambda_t N_t + qb(1 - \lambda_t)N_t}{a\lambda_t N_t + b(1 - \lambda_t)N_t} \\ &= \frac{pa\lambda_t + qb(1 - \lambda_t)}{a\lambda_t + b(1 - \lambda_t)}. \end{aligned}$$

We now provide a definition of the steady state of the population dynamics.

DEFINITION 1.  The population dynamics of an economy is in a steady state if

$$\lambda_t = \lambda_{t+1} = \lambda_{t+2} = \cdots.$$

It is easy to see that this definition implies that in steady state, the population growth rate is constant.

We next state and prove the following proposition.

PROPOSITION 2. *(1) Given a, b, p, and q, there exists a unique, stable, and nontrivial steady state. (2) At the steady state, we have*

$$\frac{\partial \lambda}{\partial a} > 0, \ \frac{\partial \lambda}{\partial b} < 0, \ \frac{\partial \lambda}{\partial p} > 0, \ \frac{\partial \lambda}{\partial q} > 0.$$

PROOF.   (1) When $\lambda_t = 1$, we have from (10)

$$\lambda_{t+1} = p < 1 = \lambda_t.$$

When $\lambda_t = 0$, we have from (10)

$$\lambda_{t+1} = q > 0 = \lambda_t.$$

From (10), we know that $\lambda_{t+1}$ is a continuous function of $\lambda_t$ on the closed interval [0, 1]. Therefore, there exists a $\lambda^*$ such that

$$\lambda_t = \lambda_{t+1} = \lambda^*.$$

Further, the locus of the dynamics of $\lambda_t$ must first pass through the 45 degree line from above (note that $\lambda_{t+1}(\lambda_t = 0) = q > 0 = \lambda_t$ and $\lambda_{t+1}(\lambda_t = 1) = p < 1 = \lambda_t$). At this intersection point—a steady state—the slope of the locus maintains

$$1 > \frac{d\lambda_{t+1}}{d\lambda_t} > 0,$$

that is, the steady state is stable.

We next prove the uniqueness of the steady state. From (10), we have that

$$(11) \qquad \frac{d\lambda_{t+1}}{d\lambda_t} = \frac{(p-q)ab}{[a\lambda_t + b(1-\lambda_t)]^2} > 0,$$

and that

$$(12) \qquad \frac{d^2\lambda_{t+1}}{d\lambda_t^2} = \frac{2(p-q)ab(b-a)}{[a\lambda_t + b(1-\lambda_t)]^3} > 0.$$

Thus, the locus of the dynamics may cross the 45 degree line again (from below), and at this possible new intersection point, we would have that $\frac{d\lambda_{t+1}}{d\lambda_t} > 1$. Since $\frac{d^2\lambda_{t+1}}{d\lambda_t^2} > 0$, we must have that for values of $\lambda_t$ that are greater than the value of $\lambda_t$ at the intersection point, $\frac{d\lambda_{t+1}}{d\lambda_t} > 1$, and hence, the locus will not cross the 45 degree line (from above) again.
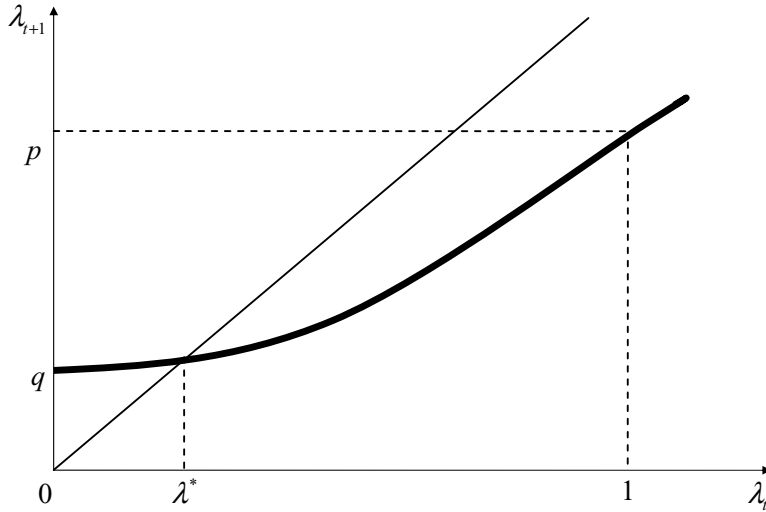
FIGURE 1

THE UNIQUE, STABLE, AND NONTRIVIAL STEADY STATE

Note, however, that $\lambda_{t+1}(\lambda_t = 1) = p < 1$. Therefore, the locus between the first intersection point and the point where $\lambda_t = 1$ must be below the 45 degree line, and hence there is no additional intersection point with the 45 degree line in this domain. We also know that $\lambda_t \leq 1$. Thus, the locus cannot pass the 45 degree line again after it passes it first and, therefore, the first intersection point defines the unique steady state of the dynamic system; $\lambda^*$ is unique.

The proof of Part (1) of Proposition 2 can be conveniently illustrated by Figure 1.

(2) At the steady state, we have from (10),

$$(13) \qquad \lambda = \frac{ap\lambda + bq(1-\lambda)}{a\lambda + b(1-\lambda)},$$

namely,

$$(14) \qquad \Omega \equiv \frac{ap\lambda + bq(1-\lambda)}{a\lambda + b(1-\lambda)} - \lambda = 0.$$

Note that

$$(15) \qquad \frac{\partial \Omega}{\partial a} = \frac{p\lambda[a\lambda + b(1-\lambda)] - \lambda[ap\lambda + bq(1-\lambda)]}{[a\lambda + b(1-\lambda)]^2}$$

$$= \frac{\lambda b(1-\lambda)(p-q)}{[a\lambda + b(1-\lambda)]^2}$$

$$> 0,$$

that

(16)
$$\frac{\partial \Omega}{\partial b} = \frac{q(1-\lambda)[a\lambda + b(1-\lambda)] - (1-\lambda)[ap\lambda + bq(1-\lambda)]}{[a\lambda + b(1-\lambda)]^2}$$
$$= -\frac{(p-q)(1-\lambda)a\lambda}{[a\lambda + b(1-\lambda)]^2}$$
$$< 0,$$

that

(17)
$$\frac{\partial \Omega}{\partial p} = \frac{a\lambda}{a\lambda + b(1-\lambda)} > 0,$$

and that

(18)
$$\frac{\partial \Omega}{\partial q} = \frac{b(1-\lambda)}{a\lambda + b(1-\lambda)} > 0.$$

Also, from the proof of the first part of the proposition, in particular from Figure 1, we know that at the steady state, the slope of the locus is less than one, that is $\frac{d\lambda_{t+1}}{d\lambda_t} < 1$. Then, from (13) and (14), we know that

(19)
$$\frac{\partial \Omega}{\partial \lambda} < 0.$$

Thus,

$$\frac{\partial \lambda}{\partial a} = -\frac{\partial \Omega}{\partial a} \Big/ \frac{\partial \Omega}{\partial \lambda} > 0,$$

$$\frac{\partial \lambda}{\partial b} = -\frac{\partial \Omega}{\partial b} \Big/ \frac{\partial \Omega}{\partial \lambda} < 0,$$

$$\frac{\partial \lambda}{\partial p} = -\frac{\partial \Omega}{\partial p} \Big/ \frac{\partial \Omega}{\partial \lambda} > 0,$$

and

$$\frac{\partial \lambda}{\partial q} = -\frac{\partial \Omega}{\partial q} \Big/ \frac{\partial \Omega}{\partial \lambda} > 0. \qquad\blacksquare$$

At the steady state, the higher the probabilities that children will be skilled, the higher the fertility of the skilled individuals; and the lower the fertility of the unskilled individuals, the higher the proportion of skilled individuals in the population.

## 4.  HETEROGENEITY AND OPTIMAL POPULATION

In this section, we will argue that a "population problem" exists because individual choices do not yield maximal social welfare. In line with the "Golden Rule"

literature on the welfare concomitants of economic growth and other dynamic analyses (*inter alia* Phelps, 1961),[9] we will focus our attention on the steady state.

Let us define the social welfare function, $W$, as an average utilitarian[10]

$$(20) \qquad W = \lambda u + (1 - \lambda)v,$$

where $u$ and $v$ are the utility of a skilled individual and that of an unskilled individual, respectively. Note that this social welfare function is equivalent to a situation that Rawls (1973) terms the "original position," which is characterized by a "veil of ignorance," an about-to-be born member of the society will obtain utility $u$ with probability $\lambda$, and utility $v$ with probability $1 - \lambda$. Thus, (20) describes the expected utility of an unborn member of the society.

We have already argued that a skilled individual enjoys a higher wage rate and is blessed with a better ability to positively affect the skill level of his children than an unskilled individual. Also, from Section 2, it is easy to see that an individual's utility increases with $w$ and $\delta$. Therefore,

$$u > v.$$

We are ready now to state and prove the following proposition.

PROPOSITION 3.  *By the standard of social optimum, a "population problem" exists in two respects:*

(1) *Skilled individuals have too few children and unskilled individuals have too many children.*
(2) *Both skilled and unskilled individuals spend too little time on children's education.*

PROOF.   Proving this proposition is equivalent to showing that given all other variables, the optimal individual choice of a particular variable is less or greater than the optimal social choice. The proof of the proposition is divided into four parts: (i) skilled individuals have too few children; (ii) unskilled individuals have too many children; (iii) skilled individuals spend too little time on children's education; (iv) unskilled individuals spend too little time on children's education. Since the underlying logic of the proofs of these four parts is quite the same, we provide next a proof of the first part; proofs of the remaining parts are relegated to the Appendix.

We seek to demonstrate that by the standard of social optimum, skilled individuals have too few children. Given all other variables, at the optimum of the individual's choice of $a$, we have that

$$(21) \qquad \frac{\partial u}{\partial a} = 0;$$

we denote the solution to this equation by $a^o$.

---

[9] Romer (2006) provides a survey of this literature.

[10] Note that an average utilitarian social welfare function is also used in much of the related literature on "optimum population," (cf. Samuelson, 1975, and Blackorby et al., 1995, 2002).

Next, from (20), we have that

$$(22) \qquad \frac{\partial W}{\partial a} = (u - v)\frac{\partial \lambda}{\partial a} + \lambda\frac{\partial u}{\partial a}.$$

From Proposition 2 we have that at the steady state, $\frac{\partial \lambda}{\partial a} > 0$. Therefore, at $a = a^o$, from (21) and (22), we have

$$(23) \qquad \frac{\partial W}{\partial a} = (u - v)\frac{\partial \lambda}{\partial a} > 0.$$

However, at the social optimum, we must have that

$$(24) \qquad \frac{\partial W}{\partial a} = 0.$$

We denote the solution to (24) by $a^*$. From (23) and (24), we know that $a^* \neq a^o$. We now seek to show that

$$a^* > a^o.$$

We do this by contradiction.

We can rewrite (20) as follows:

$$(25) \qquad W = \lambda u + (1 - \lambda)v$$
$$= \lambda(u - v) + v.$$

From (25), we can see that if $a$ decreases from $a^o$, then: (i) holding $\lambda$ constant, $(u - v)$ will decrease since $u$ achieves maximum at $a^o$ and $v$ is independent of $a$; (ii) holding $(u - v)$ constant, $\lambda$ will decrease as $\frac{\partial \lambda}{\partial a} > 0$. Since both $\lambda$ and $(u - v)$ are positive, $\lambda(u - v)$ will decrease if $a$ decreases from $a^o$. Thus, $a$ must increase from $a^o$ in order to achieve social maximum. ∎

Proposition 3 shows that individual optimal choices differ from the socially optimal choices. The intuition underlying the proposition is as follows. If parents with a high level of human capital produce more children, the proportion of skilled individuals in the next generation will be higher. Conversely, if parents with a low level of human capital produce more children, the proportion of unskilled individuals in the next generation will be higher. A "population problem" is created when, from a social welfare point of view, parents produce too few or too many children, which is likely to be the case since individual parents do not ordinarily take into account the economy-wide or social implications of their choices. Yet it is unlikely that a society will not be affected by, and hence that it would be indifferent to its skill composition and the skill levels of its members.

Proposition 3 maintains that all individuals invest too little in the education of their children. Since investment in children's education contributes to the

accumulation of human capital whereas savings by households contribute to the accumulation of physical capital, this result is akin to the finding of the neoclassical growth model that savings which are privately optimal often fall short of the level of savings that yields the golden-rule level of capital (de la Croix and Michel, 2002; Romer, 2006).[11] Thus, this article too explores the incidence of an investment that is shy of the socially optimal level of investment.

## 5.   TAXING, AND SUBSIDIZING FERTILITY

In this section, we will analyze the design of a subsidy policy aimed at increasing the fertility of skilled individuals and raising social welfare. An empirical motivation for this section is the population policy in Singapore, a country where population issues have been addressed more seriously than in a great many other countries. In an important speech at the National Day Rally on August 14, 1983, Mr. Lee Kuan Yew, the Prime Minister of Singapore at that time, expressed concern about the lopsided fertility pattern in which the better-educated women produced too few children while the lesser-educated women produced too many children, a pattern which would lower the quality of the population. Mr. Lee stated that:

> *"Therefore, we must further amend our policies, and try to reshape our demographic configuration so that our better-educated women will have more children to be adequately represented in the next generation."*[12]

In recent decades, incentives were introduced to encourage better-educated women to have more children. These incentives include "baby bonuses" for second and third births and picking up the tab for paid maternity leave for third births (see, for example, Yap, 2003).

Suppose that $s(>0)$ is the subsidy to a skilled individual for having a child—child allowance—and $t(<w)$ is a lump-sum tax. Then, in this case, an individual's budget constraint (recall (2)) becomes

$$(26) \qquad\qquad c = (1 - \xi n - nk)w - t + ns.$$

For simplicity, we will analyze only the benchmark case wherein the subsidies to the skilled are financed from the proceeds of tax levied on the skilled,[13] namely, in equilibrium

$$(27) \qquad\qquad\qquad\qquad t = ns.$$

It should be noted that when making their choices of the quantity and the quality of children individuals will treat $s$ and $t$ as parameters.

---

[11] In both the current article and in the related golden-rule literature, steady-state welfare criteria are used which essentially do not discount the future. The parents do not explicitly take their children's utility into account, but the social planner does.

[12] Lee's speech is reproduced in Appendix A of Saw (1990).

[13] Similar results can be obtained upon assuming a different tax incidence.

From (26), we can rewrite the utility function (recall (3)) as

$$(28) \qquad \ln(n) + \alpha \ln(m) + \beta \ln[(1 - \xi n - nk)w - t + ns].$$

Hence, the first order conditions for utility maximization are

$$(29) \qquad n : \frac{1}{n} = \frac{\beta[(\xi + k)w - s]}{(1 - \xi n - nk)w - t + ns},$$

$$(30) \qquad k : \frac{\alpha e_1}{e} = \frac{\beta n w}{(1 - \xi n - nk)w - t + ns}.$$

Inserting (27) into (29) and (30), we get

$$(31) \qquad \frac{1}{n} = \frac{\beta[(\xi + k)w - s]}{(1 - \xi n - nk)w},$$

$$(32) \qquad \frac{\alpha e_1}{e} = \frac{\beta n}{1 - \xi n - nk}.$$

We now have the following proposition.

PROPOSITION 4. *Child allowance will result in an increase of fertility yet in a decrease of the educational expenditure on each child if and only if the following condition is satisfied*

$$(33) \qquad (\alpha - 1)e_1^2 + ee_{11} < 0.$$

PROOF.   See the Appendix.

This proposition implies that child allowance achieves the intended impact of increasing fertility only under some technical conditions. If (33) is reversed, then child allowance will result in a decrease of fertility and in an increase of the educational expenditure on each child.

Note, in particular, that (33) will be satisfied if

$$\alpha < 1,$$

which will occur when parents care more about the quantity than about the quality of their children.

In Singapore, the government subsidizes the more educated parents in order to induce them to have more children. Proposition 4 implies that this policy will not be applicable universally but rather under some specific technical conditions.

Will the policy increase the proportion of skilled individuals in the steady state? How will it affect individual and social welfare? The answers to these questions are provided in the following proposition.

PROPOSITION 5. (1) *The sign of $\frac{d\lambda}{ds}$ is ambiguous.* (2) *u decreases as the government implements the policy of child allowance to the skilled individuals under a balanced budget for the skilled individuals.* (3) *The policy has an ambiguous impact on social welfare.*

PROOF. See the Appendix.

Proposition 5 shows that the tax and subsidy policy is riddled with implementation difficulties: It may not necessarily increase the proportion of skilled individuals in the steady state; the welfare of the skilled individuals decreases; the policy's impact on social welfare is ambiguous. Thus, in this section we have shown that a direct means aimed at solving the "population problem" may not be capable of achieving its intended purpose.

## 6. "HEAD START PROGRAM"

Up until now we have assumed that the skill trait of the *parents* is historically given, having resulted from decisions undertaken by their own parents. But what if policy intervention could impact on $\delta$? What will happen if unskilled parents were to become skilled parents? Could a government expenditure aimed at elevating $\delta$ mitigate the "population problem" and raise social welfare? Imagine, for a moment, that an exogenous shock succeeds in transforming the skill level of the unskilled from $\delta = 0$ to $\delta = 1$. Then the "population problem" dissipates away completely: The social welfare function (20) reduces to $u$, and the choice of a skilled individual coincides with the social choice ($a^o \equiv a^*$), such that whatever the individual chooses *is* the social optimum.

Our starting point in this section is a prevailing policy that, as we will seek to show, can have a desirable effect—by consequence if not by design—with regard to the "population problem." We refer to the "Head Start Program" in the United States, which "aims to improve the learning skills, social skills, and health status of poor children so that they can be in schooling on an equal footing with their more advantaged peers" (Currie and Thomas, 1995). A number of empirical studies, notably Currie and Thomas (1995), Currie (2001), and Garces et al. (2002), show that early childhood education programs have significant positive impact on the participants' educational attainments both in the short term and in the long term.[14]

To show how a "Head Start Program" can be an effective means to address the "population problem," we extend the model slightly by making the following assumption:

$$(34) \qquad\qquad \delta \equiv \delta(g), \qquad \delta'(g) > 0,$$

---

[14] Restuccia and Urrutia (2004) find that policies that target early education are much more effective in reducing poverty than policies that target college education.

where $g$ denotes the government's expenditure on improving $\delta$ for the unskilled parents.

We assume that the expenditures on the "Head Start Program" are financed by an income tax on the skilled individuals, and that the tax rate is $\gamma$. Then for skilled parents, (3) becomes

(35)    $\ln(n) + \alpha \ln(e) + \beta \ln[(1 - \xi n - kn)(1 - \gamma)w]$

$= \ln(n) + \alpha \ln(e) + \beta \ln[(1 - \xi n - kn)w] + \beta \ln(1 - \gamma).$

From (35) it is clear that an income tax proportional to an individual's earnings has no impact on the individual's choices of the quantity and the quality of children. We denote a skilled individual's optimal (private) choice of the number of children and the amount of time that he invests in the education of each of his children by $a^o$ and $x^o$, respectively. Then, assuming that the government balances its budget, we have

$$\gamma \lambda (1 - \xi a^o - x^o a^o) w_s = (1 - \lambda) g,$$

namely,

(36)    $$\gamma = \frac{1 - \lambda}{\lambda (1 - \xi a^o - a^o x^o) w_s} g.$$

From Proposition 1 and Proposition 2, we then have that

(37)    $$\frac{d\lambda}{dg} = \frac{\partial \lambda}{\partial b} \frac{db}{dg} + \frac{\partial \lambda}{\partial q} \frac{dq}{dg}$$

$$= \frac{\partial \lambda}{\partial b} \frac{\partial b}{\partial \delta} \frac{d\delta}{dg} + \frac{\partial \lambda}{\partial q} \left( \frac{\partial q}{\partial k} \frac{dk}{dg} + \frac{\partial q}{\partial \delta} \frac{d\delta}{dg} \right)$$

$$= \frac{\partial \lambda}{\partial b} \frac{\partial b}{\partial \delta} \frac{d\delta}{dg} + \frac{\partial \lambda}{\partial q} \left( \frac{\partial q}{\partial k} \frac{\partial k}{\partial \delta} \frac{d\delta}{dg} + \frac{\partial q}{\partial \delta} \frac{d\delta}{dg} \right)$$

$$= \left[ \frac{\partial \lambda}{\partial b} \frac{\partial b}{\partial \delta} + \frac{\partial \lambda}{\partial q} \left( e_1 \frac{\partial k}{\partial \delta} + e_2 \right) \right] \frac{d\delta}{dg}$$

$$> 0.$$

From (3) and (4), we can rewrite (3) as

$$v = \ln(n) + \alpha \ln[e(k, \delta)] + \beta \ln[(1 - \xi n - kn)w].$$

From the Envelop Theorem, we know that $\frac{dv}{d\delta} = \frac{\partial v}{\partial \delta}$. Therefore, we have that

$$\frac{dv}{d\delta} = \frac{\alpha e_2}{e}$$

and hence

$$(38) \qquad \frac{dv}{dg} = \frac{\alpha e_2}{e} \frac{d\delta}{dg} > 0.$$

From (35) and (36), we know that

$$(39) \qquad \frac{du}{dg} = -\frac{\beta}{1-\gamma} \left[ \frac{1-\lambda}{\lambda(1-\xi a^o - a^o x^o)w_s} - \frac{\frac{d\lambda}{dg}}{\lambda^2(1-\xi a^o - a^o x^o)w_s} g \right].$$

At $g = 0$, we have that

$$(40) \qquad \frac{du}{dg} = -\frac{\beta(1-\lambda)}{\lambda(1-\gamma)(1-\xi a^o - a^o x^o)w_s}.$$

Therefore, if $\delta'(g)$ is sufficiently large at $g = 0$, then from (25), (37), (38), and (40), we have (at $g = 0$)

$$(41)$$

$$\frac{dW}{dg} = \lambda \frac{du}{dg} - \lambda \frac{dv}{dg} + (u-v)\frac{d\lambda}{dg} + \frac{dv}{dg}$$

$$= -\frac{\beta\lambda(1-\lambda)}{\lambda(1-\gamma)(1-\xi a^o - a^o x^o)w_s} + (u-v)\left[ \frac{\partial\lambda}{\partial b}\frac{\partial b}{\partial \delta} + \frac{\partial\lambda}{\partial q}\left( e_1 \frac{\partial k}{\partial \delta} + e_2 \right) \right] \delta'(0)$$

$$+ (1-\lambda)\frac{\alpha e_2}{e}\delta'(0)$$

$$> 0.$$

However, the first order condition for the maximization of social welfare is

$$(42) \qquad \frac{dW}{dg} \leq 0$$

with strict inequality holding if $g = 0$. Thus, from (41) and (42), we know that $g = 0$ cannot satisfy the first order condition and hence it cannot be the optimal solution. Thus, we conclude by means of a proposition.

PROPOSITION 6. *The "Head Start Program" can mitigate the "population problem" and improve social welfare if $\delta'(0)$ is sufficiently large.*

As an example, if $\delta = \ln g$ or if $\delta = g^\sigma$ where $0 < \sigma < 1$, then, $\delta'(0) = \infty$.

Furthermore, when $g > 0$, if $\frac{d\delta}{dg}$ is sufficiently large then, from (37), we know that $\frac{d\lambda}{dg}$ will also be sufficiently large. In this case, from (39), we will have that $\frac{du}{dg} > 0$. Recall that $\frac{dv}{dg} > 0$ always. Thus, under these circumstances, this policy

intervention will increase the welfare of both the skilled individuals and the un-
skilled individuals in the steady state.

## 7.  HUMAN CAPITAL EXTERNALITY

In this section we add to the preceding analyses by incorporating the externality
effect of average human capital. Specifically, we assume that the wage rate of the
skilled individuals in a given period is positively related to the average level of
human capital in the economy in that period. Namely, we assume that

$$(43) \qquad\qquad w_s = w_s(\lambda), \qquad \frac{dw_s}{d\lambda} > 0.$$

In such a setting, we explore a scheme wherein the government uses pecuniary
incentives to induce unskilled individuals to reduce fertility, which mitigates the
"population problem." We show that under reasonable conditions, the scheme can
improve the welfare of the unskilled individuals and the net income of the skilled
individuals in every generation. The scheme is not divorced from concrete em-
pirical settings. For example, in the 1980s, the government of Singapore offered
grants of 10,000 Singaporean dollars to little-educated women who agreed to be
sterilized after the birth of their second child conditional on the following require-
ments (along with several secondary requirements) being met: (i) Neither parent
should have any General Certificate of Education (GCE) Ordinary Level passes;
(ii) neither parent should earn more than 750 Singaporean dollars per month (see
Saw, 1999).

The scheme is described as follows:

Period 1:  The government borrows to pay the unskilled to "bribe" or induce
   them to reduce their fertility.
Period 2:  The skilled of generation 2 pay for the government-incurred debt of
   the last generation; the government borrows to pay the "new" unskilled to
   "bribe" or induce them to reduce their fertility.
Period 3:  The skilled of generation 3 pay for the government debt of the last
   generation; the government borrows to pay the unskilled to "bribe" or induce
   them to reduce fertility.
And so on.

The basic logic of this scheme is straightforward: If unskilled individuals in the
current generation reduce fertility, the proportion of skilled individuals in the next
generation will be greater, which increases the wage rate of skilled individuals of
the next generation under the assumption in (43). Thus, if the externality effect of
average human capital is large enough, the cost of inducing unskilled individuals
to have a lower level of fertility will be less than the resulting gain for the skilled
individuals in the next generation.

We now analyze the scheme in detail.

Under the status quo, an unskilled individual's optimal (private) choice of the
number of children and the amount of time that he invests in the education of
each of his children are denoted by $b^o$ and $y^o$, respectively. We assume that the tax

levied on a skilled individual is proportional to his wage rate. Then, as illustrated in Section 6 (Equation (35)), the scheme will not affect a skilled individual's optimal choice of the number of children and the amount of time that he invests in the education of each of his children which, to recall, are denoted by $a^o$ and $x^o$, respectively.

(1) Generation 1: The government increases the wage rate of the unskilled from $w_u$ to $(1 + \tau)w_u$. At the same time, the government announces that an unskilled individual will receive the benefit if and only if he chooses his fertility level to be $b^*$, which is the optimal level of $b$ that maximizes social welfare. From Proposition 3, it follows that $b^* < b^o$. If an unskilled individual chooses $b^*$ number of children, he will choose a new value of $k$, the amount of time that he invests in the education of each of his children, that maximizes his utility. We denote this new level of $k$ by $y^n$.

We next examine under what condition an unskilled individual will accept the government's offer. From the preceding analyses we can see that an individual's utility can be written as a function of $n$ and $k$, and we denote this function by $V(n, k)$. Under the status quo, an unskilled individual's utility arising from the individual's ability to optimally choose the quantity and the quality of his children (that is, $n = b^o$ and $k = y^o$) is

$$(44) \qquad\qquad\qquad V(b^o, y^o).$$

If the unskilled individual accepts the government's offer, he will choose $n = b^*$ and $k = y^n$. Recalling (3), the unskilled individual's utility is then

$$(45) \qquad \ln(n) + \alpha \ln(e) + \beta \ln[(1 - \xi n - kn)(1 + \tau)w_u]$$
$$= V(b^*, y^n) + \beta \ln(1 + \tau).$$

An unskilled individual will therefore accept the government's offer if and only if

$$(46) \qquad\qquad V(b^*, y^n) + \beta \ln(1 + \tau) \geq V(b^o, y^o).$$

Namely, if and only if

$$(47) \qquad\qquad \tau \geq \exp\left[\frac{V(b^o, y^o) - V(b^*, y^n)}{\beta}\right] - 1.$$

We assume that the government chooses the following $\tau$

$$(48) \qquad\qquad \tau = \exp\left[\frac{V(b^o, y^o) - V(b^*, y^n)}{\beta}\right] - 1 + \varepsilon,$$

where $\varepsilon$ is a small positive number. Thus, implementation of the scheme will increase the welfare of the unskilled individuals. Then, the debt of the government is

(49)

$$D_1 \equiv (1 - \lambda_1)N_1(1 - \xi b^* - y^n b^*)\tau w_u$$

$$= (1 - \lambda_1)N_1(1 - \xi b^* - y^n b^*)w_u \left\{ \exp\left[\frac{V(b^o, y^o) - V(b^*, y^n)}{\beta}\right] - 1 + \varepsilon \right\},$$

where $\lambda_1$ and $N_1$ denote the proportion of skilled individuals in Generation 1 and the total population in Generation 1, respectively.

(2) Generation 2: We denote by $\lambda_2^*$ the proportion of skilled individuals in the current generation (that is, in Generation 2) when the unskilled individuals of the last generation accepted the government's offer, and we denote by $\lambda_2$ the proportion of the skilled individuals in the current generation (that is, Generation 2) under the status quo. We then have the following proposition.

PROPOSITION 7. (1): $\lambda_2^* > \lambda_2$. (2) *If the externality effect of average human capital is large enough, the scheme delineated in this section can improve the welfare of the unskilled individuals and the net income of the skilled individuals in every generation.*

PROOF. See the Appendix.

## 8. ALTERNATIVE SOCIAL WELFARE FUNCTIONS

This section offers an inquiry into the validity of our main analytical results under alternative social welfare functions. A brief summary of our findings is in order. First, if the social welfare function is Rawlsian, which is defined as the welfare of the worst-off individual in the society, then we show that there will be no "population problem." However, as is commonly recognized, the Rawlsian social welfare function has limited applicability and appeal.[15] Second, if we let the social welfare function to be a "generalized" average utilitarian, as defined by Blackorby et al. (1995, 2002), we get results that are qualitatively the same as Proposition 3. Thus, we infer that the results obtained in Section 4 are fairly robust. Finally, we consider the case wherein the social welfare function is a function both of the population growth rate and of the individuals' average utility, in the spirit of total utilitarianism. We show that in such a case, the results obtained in Section 4 will continue to hold under reasonable additional assumptions.

8.1. *A Rawlsian Social Welfare Function.* The Rawlsian social welfare function is defined as the welfare of the worst-off individual in society (see, for example, Nerlove et al., 1988). In our model, there are only two types of individuals. From Section 4, we know that

$$u > v.$$

---

[15] For example, in the political economy literature, it is usually assumed that political outcomes are determined by "median voters" (cf. Merlo, 2006).

Thus, in our model the Rawlsian social welfare function is simply the unskilled individual's utility function, namely,

$$(50) \qquad\qquad W = v.$$

This implies that the maximizing of social welfare collapses to the maximizing of the utility of an unskilled individual. Therefore, the private choices of an unskilled individual coincide with the social optimum. Also, a skilled individual's choices obviously do not affect an unskilled individual's welfare (the wage rate of an unskilled individual, and for that matter everything else, is independent of a skilled individual's choices), which implies that a skilled individual's private choices (which do not affect $v$) are also consistent with the social optimum.

In summary, we have the following proposition.

PROPOSITION 8. *Suppose that the social welfare function is Rawlsian. Then, by the standard of social optimum, there is no "population problem."*

Proposition 8 is derived under the assumption of no government intervention such as tax and subsidy policies. When the government does play a role, however, then under the Rawlsian social welfare function, social welfare will be improved, for example by the "Head Start Program" elucidated in Section 6, or by a policy of taxing the skilled individuals and transferring the proceeds to the unskilled individuals.

8.2. *A "Generalized" Average Utilitarian Social Welfare Function.* As, for example, in Blackorby et al. (1995, 2002), we now define the social welfare function to be "generalized" average utilitarian, namely,

$$(51) \qquad\qquad W = (1 - \rho)\lambda u + \rho(1 - \lambda)v,$$

where, to recall, $\lambda$ is the steady-state proportion of skilled individuals in the population, and where $\rho$ is a coefficient, $0 < \rho < 1$; the greater $\rho$, the greater the weight accorded to unskilled individuals in the social welfare calculus. We make the following assumption:

$$(52) \qquad\qquad (1 - \rho)u > \rho v,$$

namely,

$$(53) \qquad\qquad \rho < \frac{u}{u + v}.$$

Recalling that $u > v$, we know that $\frac{u}{u+v} > \frac{1}{2}$. Thus, we may well have $\rho > \frac{1}{2}$, which still satisfies (53). Condition (53) means that while more weight can be accorded

to unskilled individuals than to skilled individuals ($\rho > \frac{1}{2}$), the weight is bounded by the value $\frac{u}{u+v}$. Based on this assumption, we have the following proposition.

PROPOSITION 9.  *Suppose that the social welfare function is defined in (51), and suppose that the condition (53) is satisfied. Then, by the standard of social optimum, a "population problem" exists in two respects:*

(1) *Skilled individuals have too few children and unskilled individuals have too many children.*
(2) *Both skilled and unskilled individuals spend too little time on children's education.*

PROOF.   See the Appendix.

Proposition 9 generalizes Proposition 3 and implies that the results obtained in Section 4 are fairly robust.

8.3. *Social Welfare as a Function both of the Population Growth Rate and of Individuals' Average Utility.*    We consider a setting wherein the social welfare function is a function of both the population growth rate and of the individuals' average utility, in the spirit of total utilitarianism.[16] From the analysis in Section 3, we know that the rate of population growth in each period at the steady state, which we denote by $\pi$, is

$$(54) \qquad \pi = \frac{a\lambda N_t + b(1-\lambda)N_t}{N_t} - 1$$
$$= a\lambda + b(1-\lambda) - 1.$$

Then, the social welfare function is defined as follows:

$$(55) \qquad\qquad Z = S(W, \pi),$$

where $Z$ denotes social welfare and $W$ is defined in (20).

For simplicity, in this subsection we only consider whether and how individuals' choices of fertility may deviate from social optima. The first order conditions for the maximization of $Z$ with respect to $a$ and $b$ are, respectively,

$$(56) \qquad\qquad \frac{dZ}{da} = \frac{\partial Z}{\partial W}\frac{dW}{da} + \frac{\partial Z}{\partial \pi}\frac{d\pi}{da} = 0$$

---

[16] According to total utilitarianism, the social welfare is the sum of the utilities of all the individuals (Nerlove et al., 1986a; Razin and Sadka, 1995). In an infinite-period framework, the utility of future generations is incorporated with a social discount factor so that the function is well defined (Gigliotti, 1983). The key elements in total utilitarianism are the total number of individuals, and the utility of every individual. Since the rate of population growth determines the total number of individuals, our formulation here is in the spirit of total utilitarianism. Consequently, our formulation significantly simplifies the algebra without qualitatively changing the spirit or the implications of total utilitarianism.

and

$$(57) \qquad \frac{dZ}{db} = \frac{\partial Z}{\partial W}\frac{dW}{db} + \frac{\partial Z}{\partial \pi}\frac{d\pi}{db} = 0.$$

We assume that the second order conditions of $Z$ with respect to $a$ and $b$ are negative: $\frac{dZ}{da}$ is a decreasing function of $a$, and $\frac{dZ}{db}$ is a decreasing function of $b$. Thus, obviously, we have the following proposition.

PROPOSITION 10. *Suppose that the social welfare function is defined in (55). Then, by the standard of social optimum, a "population problem" exists in the following respects:*

(1) *Skilled individuals have too few children if and only if when $a = a^o$,*

$$(58) \qquad \frac{\partial Z}{\partial W}\frac{dW}{da} + \frac{\partial Z}{\partial \pi}\frac{d\pi}{da} > 0.$$

(2) *Unskilled individuals have too many children if and only if when $b = b^o$,*

$$(59) \qquad \frac{\partial Z}{\partial W}\frac{dW}{db} + \frac{\partial Z}{\partial \pi}\frac{d\pi}{db} < 0.$$

This proposition implies that the results obtained in Section 4 continue to hold under reasonable assumptions in this new setting. Moreover, note that

$$(60) \qquad \frac{d\pi}{da} = \lambda - (b - a)\frac{\partial \lambda}{\partial a}.$$

Equation (60) implies that we may have either $\frac{d\pi}{da} > 0$ or $\frac{d\pi}{da} < 0$. The intuition is as follows. On one hand, if skilled individuals have more children (that is, if $a$ increases), this will contribute to an increase in population growth. On the other hand, when skilled individuals have more children, there will be a higher proportion of skilled individuals in the next generation. Since a skilled individual has fewer children than an unskilled individual, the result will be a decrease in population growth in the long run. Thus, in the steady state, the impact of an increase in the fertility rate of skilled individuals has an ambiguous effect on the population growth rate.

For unskilled individuals we have

$$(61) \qquad \frac{d\pi}{db} = 1 - \lambda - (b - a)\frac{\partial \lambda}{\partial b} > 0.$$

The inequality in (61) implies that if an unskilled individual has more children (that is, if $b$ increases), the population growth rate in the steady state will surely increase.

When the society has a strong desire for a faster population growth, $\frac{\partial Z}{\partial \pi}$ is large. In this case, from Proposition 10, the following comments ensue. First, since $\frac{d\pi}{db} > 0$,

if $\frac{\partial Z}{\partial \pi}$ is sufficiently large (such that (59) is not satisfied), then there will no longer be a "population problem" of unskilled workers having too many children. This is so because an unskilled worker's having more children contributes to population growth in the very society that strongly desires a faster rate of population growth, although the average utility of the future generation will subsequently be reduced. Second, if $\frac{d\pi}{da} > 0$, from the logic of the proof of Proposition 3, we can see that the population problem will exist for skilled workers having too few children. Third, if $\frac{d\pi}{da} < 0$ and if $\frac{d\pi}{da}$ is sufficiently small so that (58) is not satisfied, then the population problem will not exist for skilled workers having too few children. The intuition is that if an increase in skilled individuals' fertility rate decreases the population growth rate and when the society strongly desires rapid population growth, then the impact of the said desire will outweigh the impact of an increase in skilled individuals' fertility rate on average utility. Consequently, social welfare will not increase when skilled individuals have more children.

Thus, in summary, the population problem analyzed in this paper will continue to exist if a society's desire for a fast population growth is not too strong (such that (58) and (59) are satisfied). In other words, our main analytical results will hold in this alternative setting under additional reasonable conditions. Notably, in developing countries a slow population growth is not a major concern, or even not a concern at all. Then, $\frac{\partial Z}{\partial \pi}$ is small so that (58) and (59) are likely to be satisfied, and the population problem identified in Section 4 continues to exist.

## 9.  CONCLUSION

Across different countries, fertility rates are low in rich countries and high in poor countries, and in a given country, poorer households have more children than richer households. Indeed, the negative correlation between the level of fertility and the level of income is one of the most robust empirical findings in demography and in economics. In particular, due to the strong association between poverty and high fertility, the "population problem" has for long been of concern to researchers in various social science disciplines and, of course, to policy makers.

This paper investigates the "population problem" by introducing heterogeneity into a Becker-Lewis model and extending the model to a dynamic framework. In line with the "Golden Rule" literature on the welfare analysis of economic growth, we focus our analysis on the steady state of a dynamic system. In our model, the steady state is defined as the situation in which the proportion of skilled individuals becomes constant in the dynamic system. Our analysis points out that from a social welfare point of view, parents either produce too few or too many children *and* invest too little in the education of their children, which is likely to be the case since they ordinarily ignore the impact of their fertility choices and outcomes on the proportion of skilled individuals in the next generation(s). Thus, a "population problem" exists. While the "population problem" is exacerbated if we assume that the average level of human capital in a population affects productivity, our results

do not hinge on us resorting to human capital externalities in production.[17] An extension of the model indicates that our results are fairly robust under alternative social welfare functions.

We investigate policy tools geared at correcting the "population problem." We inquire how child allowances and other tax-subsidy policies can be harnessed to correct the "population problem" and enhance welfare. (Singapore's subsidization of the fertility of the educated is a case in point.) We also show that early childhood education programs are effective in mitigating the "population problem." In addition, in the presence of an externality effect of average human capital, we explore a scheme that can improve the welfare of the unskilled individuals and the net income of the skilled individuals in every generation.

We have sought to add to the received welfare and policy analysis of the population problem. For simplicity, by adopting a framework of a small open economy, we have abstracted from general equilibrium effects on wages and interest rates. In future research, the model could well be extended to examine the possibility that an increase in the number of skilled workers may have ambiguous effects on welfare, since the increase may affect the wage rate. In such a setting, we would need to consider both the impact of increased labor supply on the wage rates and the channel through which an increase in skilled labor may affect the technological infrastructure of the economy (Acemoglu, 1998). Also, in an open economy world that consists of both poor and rich countries, as studied by Razin et al. (2002) and de la Croix and Gosseries (2007), we may want to consider the manner in which the population problem is related to international migration.

## APPENDIX: PROOFS

PROOF OF PROPOSITION 1.    From (5), we get

$$(A.1) \qquad\qquad n = \frac{1}{(1 + \beta)(\xi + k)}.$$

From (A.1), we have

$$(A.2) \qquad\qquad \frac{dn}{dk} = -\frac{1}{(1 + \beta)(\xi + k)^2} < 0.$$

Recalling that $\frac{\partial m}{\partial k} > 0$, we have that this result implies that there is a trade-off between the quantity and the quality of children.

---

[17] Since our analysis focuses on the steady state, it does not pertain to the issue of aging, which results from a sharp fertility decline trekking a sharp fertility increase. If the global economy were to be considered as a single economy, then the "population problem" could be conceived as one wherein fertility rates in rich countries are too low, while fertility rates in poor countries are too high.

From (5) and (6), we get

(A.3)
$$\frac{\alpha e_1}{e} = \frac{1}{\xi + k},$$

or that

(A.4)
$$\alpha e_1(\xi + k) - e = 0.$$

Totally differentiating (A.4) with respect to $k$ and $\delta$ and rearranging, we get

(A.5)
$$\frac{dk}{d\delta} = -\frac{\alpha(\xi + k)e_{12} - e_2}{\alpha(\xi + k)e_{11} + \alpha e_1 - e_1}.$$

Note that drawing on (5), the first order condition of (3) with respect to $k$ can be reduced to (A.4). Then, from (A.4), it is easy to verify that the denominator in (A.5) is just the second order condition of (3) with respect to $k$, which must be negative at the optimum. Thus, if

(A.6)
$$\alpha(\xi + k)e_{12} - e_2 > 0,$$

we will have that

$$\frac{dk}{d\delta} > 0.$$

Then, if $\frac{dk}{d\delta} > 0$, then recalling (A.2), we have

(A.7)
$$\frac{dn}{d\delta} = \frac{dn}{dk}\frac{dk}{d\delta} < 0.$$

Meanwhile, when $\frac{dk}{d\delta} > 0$, we have from (4) that

(A.8)
$$\frac{dm}{d\delta} = \frac{\partial e}{\partial k}\frac{dk}{d\delta} + \frac{\partial e}{\partial \delta} > 0. \qquad \blacksquare$$

PROOF OF PARTS (ii), (iii), AND (iv) OF PROPOSITION 3.   The proof of part (i) was provided in the text. In parts (ii), (iii), and (iv) we show that unskilled individuals have too many children, that skilled individuals spend too little time on children's education, and that unskilled individuals spend too little time on children's education, respectively.

(ii) Given other variables, at the optimum of individual choice of $b$, we have that

(A.9)
$$\frac{\partial v}{\partial b} = 0.$$

We denote the solution to (A.9) by $b^o$.

Next, from (20), we have that

$$(A.10) \qquad \frac{\partial W}{\partial b} = u\frac{\partial \lambda}{\partial b} - v\frac{\partial \lambda}{\partial b} + (1-\lambda)\frac{\partial v}{\partial b}$$

$$= (1-\lambda)\frac{\partial v}{\partial b} + (u-v)\frac{\partial \lambda}{\partial b}.$$

From Proposition 2, we have that at the steady state $\frac{\partial \lambda}{\partial b} < 0$. Therefore, at $b = b^o$, from (A.9) and (A.10), we have

$$(A.11) \qquad \frac{\partial W}{\partial b} = (u-v)\frac{\partial \lambda}{\partial b} < 0.$$

However, at the social optimum, we must have that

$$(A.12) \qquad \frac{\partial W}{\partial b} = 0.$$

We denote the solution to (A.12) by $b^*$. From (A.11) and (A.12), we know that $b^* \neq b^o$. We now seek to show that

$$b^* < b^o.$$

We do this by contradiction.

We can rewrite (20) as

$$(A.13) \qquad \begin{aligned} W &= \lambda u + (1-\lambda)v \\ &= u - (1-\lambda)(u-v). \end{aligned}$$

From (A.13), we can see that if $b$ increases from $b^o$, then: (1) holding $(1-\lambda)$ constant, $(u-v)$ will increase since $v$ achieves maximum at $b^o$ and $u$ is independent of $b$; (2) holding $(u-v)$ constant, $(1-\lambda)$ will increase as $\frac{\partial \lambda}{\partial b} < 0$. Since both $(1-\lambda)$ and $(u-v)$ are positive, $-(1-\lambda)(u-v)$ will decrease if $b$ increases from $b^o$. Thus, $b$ must decrease from $b^o$ in order to achieve social maximum.

(iii) Let $x$ be the level of $k$ that is optimally chosen by a skilled individual. Thus,

$$p = e(x, 1).$$

Then, given other variables, at the optimum of the individual choice of $x$, we have that

$$(A.14) \qquad \frac{\partial u}{\partial x} = 0.$$

We denote the solution to (A.14) by $x^o$.

Next, from (20), we have

(A.15)
$$\frac{\partial W}{\partial x} = (u - v)\frac{\partial \lambda}{\partial p}\frac{\partial p}{\partial x} + \lambda\frac{\partial u}{\partial x}$$

$$= e_1(x, 1)(u - v)\frac{\partial \lambda}{\partial p} + \lambda\frac{\partial u}{\partial x}.$$

From Proposition 2, we have that at the steady state $\frac{\partial \lambda}{\partial p} > 0$. Therefore, at $x = x^o$, from (A.14) and (A.15), we have

(A.16)
$$\frac{\partial W}{\partial x} = e_1(x, 1)(u - v)\frac{\partial \lambda}{\partial p} > 0.$$

However, at the social optimum, we must have that

(A.17)
$$\frac{\partial W}{\partial x} = 0.$$

We denote the solution to (A.17) by $x^*$. From (A.16) and (A.17), we know that $x^* \neq x^o$. We now seek to show that

$$x^o < x^*.$$

We do this by contradiction.

From (25), we can see that if $x$ decreases from $x^o$, then: (1) holding $\lambda$ constant, $(u - v)$ will decrease since $u$ achieves maximum at $x^o$ and $v$ is independent of $x$; (2) holding $(u - v)$ constant, $\lambda$ will decrease as $\frac{\partial \lambda}{\partial p} > 0$ and $\frac{\partial p}{\partial x} > 0$. Since both $\lambda$ and $(u - v)$ are positive, $\lambda(u - v)$ will decrease if $x$ decreases from $x^o$. Thus, $x$ must increase from $x^o$ in order to achieve social maximum.

(iv) Let $y$ be the level of $k$ that is optimally chosen by an unskilled individual. Thus,

$$q = e(y, 0).$$

Then, given other variables, at the optimum of the individual choice of $y$, we have that

(A.18)
$$\frac{\partial v}{\partial y} = 0.$$

We denote the solution to (A.18) by $y^o$.

Next, from (20), we have that

(A.19)
$$\frac{\partial W}{\partial y} = u\frac{\partial \lambda}{\partial q}\frac{\partial q}{\partial y} - v\frac{\partial \lambda}{\partial q}\frac{\partial q}{\partial y} + (1 - \lambda)\frac{\partial v}{\partial y}$$

$$= (1 - \lambda)\frac{\partial v}{\partial y} + e_1(y, 0)(u - v)\frac{\partial \lambda}{\partial q}.$$

From Proposition 2, we have that at the steady state $\frac{\partial \lambda}{\partial q} > 0$. Therefore, at $y = y^o$, from (A.18) and (A.19), we have

(A.20) $$\frac{\partial W}{\partial y} = e_1(y, 0)(u - v)\frac{\partial \lambda}{\partial q} > 0.$$

However, at the social optimum, we must have that

(A.21) $$\frac{\partial W}{\partial y} = 0.$$

We denote the solution to (A.21) by $y^*$. From (A.20) and (A.21), we know that $y^* \neq y^o$. We now seek to show that

$$y^o < y^*.$$

We do this by contradiction.

From (A.13), we can see that if $y$ decreases from $y^o$, then: (1) holding $(1 - \lambda)$ constant, $(u - v)$ will increase since $v$ achieves maximum at $y^o$ and $u$ is independent of $y$; (2) holding $(u - v)$ constant, $(1 - \lambda)$ will increase as $\frac{\partial \lambda}{\partial y} > 0$. Since both $(1 - \lambda)$ and $(u - v)$ are positive, $-(1 - \lambda)(u - v)$ will decrease if $y$ decreases from $y^o$. Thus, $y$ must increase from $y^o$ in order to achieve social maximum. ∎

PROOF OF PROPOSITION 4.  From (31) and (32), we get

(A.22) $$\frac{\alpha e_1}{ew} = \frac{1}{(\xi + k)w - s}$$

or that

(A.23) $$s = (\xi + k)w - \frac{ew}{\alpha e_1}.$$

Thus, recalling that $m = e(k, \delta)$,

(A.24) $$\frac{ds}{dk} = \frac{d[(\xi + k)w - \frac{ew}{\alpha e_1}]}{dk}$$
$$= w\left(1 - \frac{e_1^2 - ee_{11}}{\alpha e_1^2}\right)$$
$$= \frac{w[(\alpha - 1)e_1^2 + ee_{11}]}{\alpha e_1^2}$$

or

(A.25) $$\frac{dk}{ds} = \frac{\alpha e_1^2}{w[(\alpha - 1)e_1^2 + ee_{11}]}.$$

From (31), we get

(A.26) $$w = n[(1 + \beta)(\xi + k)w - \beta s].$$

Totally differentiating (A.26) with respect to $k, n$, and $s$, we get

(A.27) $$0 = [(1 + \beta)(\xi + k)w - \beta s]\, dn + n[(1 + \beta)w\, dk - \beta\, ds].$$

From (A.25), we get

(A.28) $$dk = \frac{\alpha e_1^2}{w[(\alpha - 1)e_1^2 + ee_{11}]}\, ds.$$

Inserting (A.28) into (A.27) yields

(A.29) $$0 = [(1 + \beta)(\xi + k)w - \beta s]\, dn + n\left[\frac{(1 + \beta)w\alpha e_1^2}{w[(\alpha - 1)e_1^2 + ee_{11}]}\, ds - \beta\, ds\right].$$

Rearranging (A.29), we get

(A.30) $$\frac{dn}{ds} = \frac{n[\beta ee_{11} - (\alpha + \beta)e_1^2]}{[(\alpha - 1)e_1^2 + ee_{11}][(1 + \beta)(\xi + k)w - \beta s]}.$$

Note that from (A.26), we must have

$$[(1 + \beta)(\xi + k)w - \beta s] > 0.$$

Meanwhile, clearly,

$$\beta ee_{11} - (\alpha + \beta)e_1^2 < 0.$$

Hence, if

(A.31) $$(\alpha - 1)e_1^2 + ee_{11} < 0$$

we will have from (A.30) that

$$\frac{dn}{ds} > 0,$$

and from (A.25) that

$$\frac{dk}{ds} < 0. \qquad\qquad\blacksquare$$

PROOF OF PROPOSITION 5.   First, we note that

(A.32)
$$\frac{d\lambda}{ds} = \frac{\partial\lambda}{\partial a}\frac{\partial a}{\partial s} + \frac{\partial\lambda}{\partial p}\frac{\partial p}{\partial k}\frac{\partial k}{\partial s}$$
$$= \frac{\partial\lambda}{\partial a}\frac{\partial a}{\partial s} + e_1\frac{\partial\lambda}{\partial p}\frac{\partial k}{\partial s}.$$

To simplify, we define

$$\Gamma \equiv (1+\beta)(\xi+k)w - \beta s,$$
$$\Delta \equiv (\alpha-1)e_1^2 + ee_{11}.$$

From the preceding analysis, we know that $\Gamma > 0$ and $\Delta < 0$.
  We denote

$$\Pi \equiv 2\lambda(b-a) - b + ap - bq.$$

From the analysis in Section 3, it can be derived that

(A.33)
$$\frac{\partial\lambda}{\partial a} = -\frac{\partial\Omega}{\partial a}\bigg/\frac{\partial\Omega}{\partial\lambda} = -\frac{\lambda(p-\lambda)}{\Pi}$$

and that

(A.34)
$$\frac{\partial\lambda}{\partial p} = -\frac{\partial\Omega}{\partial p}\bigg/\frac{\partial\Omega}{\partial\lambda} = -\frac{a\lambda}{\Pi}.$$

Thus, noting that in this case $a \equiv n$, we have

(A.35)
$$\frac{d\lambda}{ds} = \frac{\partial\lambda}{\partial a}\frac{da}{ds} + e_1\frac{\partial\lambda}{\partial p}\frac{dk}{ds}$$
$$= -\frac{\lambda(p-\lambda)}{\Pi}\frac{n[\beta ee_{11} - (\alpha+\beta)e_1^2]}{\Delta\Gamma} + e_1\left(-\frac{a\lambda}{\Pi}\right)\frac{\alpha e_1^2}{w\Delta}$$
$$= -\frac{a\lambda}{w\Delta\Gamma\Pi}[w(p-\lambda)[\beta ee_{11} - (\alpha+\beta)e_1^2] + \alpha e_1^3\Gamma].$$

Note that

$$w(p-\lambda)[\beta ee_{11} - (\alpha+\beta)e_1^2] < 0;$$
$$\alpha e_1^3\Gamma > 0.$$

Therefore, the sign of $\frac{d\lambda}{ds}$ is ambiguous.
  Next, it should be noted that the child allowance policy *reduces u*. Note that for skilled individuals, when $n = a^o$ and $k = y^o$, $U$ reaches the maximum. However,

under the policy, individuals' choices of the quantity and the quality of children are determined by the first order conditions (31) and (32). Note that (32) is the same as (6), yielding the same expression of $k$, as a function of $n$, as the expression in the absence of the policy. Therefore, the optimal solutions $n(a)$ and $k(x)$ can be obtained by inserting $k$ as a function of $n$ into (31) and (5), respectively. However, since $s > 0$, it is easy to see that (31) differs from (5), hence the solutions that satisfy (31) cannot satisfy the first order condition (5), which implies that these solutions cannot maximize $U$.

Finally, recall that $W = \lambda(u - v) + v$. Because the child allowance policy with a balanced budget for the skilled does not necessarily increase $\lambda$, and it decreases $u$, it clearly has an ambiguous impact on social welfare.                           ∎

PROOF OF PROPOSITION 7.   (1) In the case that $n = b^*$, the choice variable of an unskilled individual is only $k$ (since he treats $n$ as a parameter). Then, the first order condition for the maximization of (3) is

$$k : \frac{\alpha e_1}{e} = \frac{\beta n}{1 - \xi n - kn},$$

namely,

$$(A.36) \qquad\qquad \alpha e_1(1 - \xi n - kn) = \beta n e.$$

Totally differentiating (A.36) with respect to $n$ and $k$ and rearranging, we get

$$(A.37) \qquad \frac{dk}{dn} = \frac{\alpha e_1(\xi + k) + \beta e}{\alpha(1 - \xi n - kn)e_{11} - \alpha n e_1 - \beta n e_1} < 0.$$

Since $b^* < b^o$, (A.37) implies that $y^n > y^o$.

Next, from (10), we have that

$$(A.38) \qquad \frac{d\lambda_{t+1}}{db} = -\frac{(p - q)(1 - \lambda)a\lambda_t}{[a\lambda_t + b(1 - \lambda_t)]^2} < 0,$$

From (10) and (4), we have that

$$(A.39) \qquad \frac{d\lambda_{t+1}}{dk} = \frac{d\lambda_{t+1}}{dq}\frac{dq}{dk} = \frac{b(1 - \lambda_t)e_1}{a\lambda_t + b(1 - \lambda_t)} > 0.$$

Since $b^* < b^o$ and $y^n > y^o$, from (A.38) and (A.39), we have that

$$(A.40) \qquad\qquad\qquad \lambda_2^* > \lambda_2.$$

(2) From (43) and (A.40), we have that

$$(A.41) \qquad\qquad\qquad w_s(\lambda_2^*) > w_s(\lambda_2).$$

Therefore, the net gain in the wage rate of a skilled individual in Generation 2 is

$$(A.42) \qquad\qquad w_s(\lambda_2^*) - w_s(\lambda_2),$$

and the total gain of earnings for skilled individuals in Generation 2 is

$$(A.43) \qquad\quad G_2 \equiv \lambda_2^* N_2 (1 - \xi a^o - x^o a^o)[w_s(\lambda_2^*) - w_s(\lambda_2)],$$

where $N_2$ denotes the total population in Generation 2. Clearly,

$$(A.44) \qquad\qquad N_2 = [a^o \lambda_1 + b^*(1 - \lambda_1)]N_1.$$

If

$$(A.45) \qquad\qquad G_2 \geq (1 + r)D_1,$$

where $r$ is the interest rate, then the gain to the total income of the skilled individuals in Generation 2 will be higher than the government's cost of implementing the scheme in Generation 1. From (49), (A.43), and (A.44), we can rewrite (A.45) as

$$(A.46) \qquad \lambda_2^*[a^o \lambda_1 + b^*(1 - \lambda_1)]N_1(1 - \xi a^o - x^o a^o)[w_s(\lambda_2^*) - w_s(\lambda_2)]$$
$$\geq (1+r)(1-\lambda_1)N_1(1 - \xi b^* - y^n b^*)w_u$$
$$\times \left\{ \exp\left[\frac{V(b^o, y^o) - V(b^*, y^n)}{\beta}\right] - 1 + \varepsilon \right\},$$

namely, as

$$(A.47) \qquad w_s(\lambda_2^*) - w_s(\lambda_2) \geq \frac{(1+r)(1-\lambda_1)(1 - \xi b^* - y^n b^*)w_u}{\lambda_2^*[a^o \lambda_1 + b^*(1 - \lambda_1)](1 - \xi a^o - x^o a^o)}$$
$$\times \left\{ \exp\left[\frac{V(b^o, y^o) - V(b^*, y^n)}{\beta}\right] - 1 + \varepsilon \right\}.$$

If the externality effect of average human capital is large enough, $w_s(\lambda_2^*) - w_s(\lambda_2)$ will be large. In this case, (A.47) and hence (A.45) will be satisfied, which means that the gain of the total income of the skilled individuals in Generation 2 is greater than the cost of inducing the unskilled individuals to reduce fertility in Generation 1.

By similar logic, the process can be repeated for subsequent generations. Thus, under reasonable conditions, the scheme can improve the welfare of the unskilled individuals and the net income of the skilled individuals in every generation. ∎

PROOF OF PROPOSITION 9. The proof of this proposition is quite similar to the proof of Proposition 3. As in the case of the proof of Proposition 3, the proof of Proposition 9 is divided into four parts with the same underlying logic:

(i) skilled individuals have too few children; (ii) unskilled individuals have too many children; (iii) skilled individuals spend too little time on children's education; (iv) unskilled individuals spend too little time on children's education. Thus, for the sake of brevity, we only demonstrate that by the standard of social optimum, skilled individuals have too few children.

Given all other variables, at the optimum of the individual's choice of $a$, we have

$$(A.48) \qquad \frac{\partial u}{\partial a} = 0.$$

As noted in Section 4, we denote the solution to this equation by $a^o$.

Next, from (51), we have that

$$(A.49) \qquad \frac{\partial W}{\partial a} = [(1 - \rho)u - \rho v]\frac{\partial \lambda}{\partial a} + (1 - \rho)\lambda \frac{\partial u}{\partial a}.$$

From Proposition 2, we have that at the steady state, $\frac{\partial \lambda}{\partial a} > 0$. Therefore, at $a = a^o$, from (53), (A.48), and (A.49), we have

$$(A.50) \qquad \frac{\partial W}{\partial a} = [(1 - \rho)u - \rho v]\frac{\partial \lambda}{\partial a} > 0.$$

However, at the social optimum, we must have that

$$(A.51) \qquad \frac{\partial W}{\partial a} = 0.$$

We denote the solution to (A.51) by $a^{**}$. From (A.50) and (A.51), we know that $a^{**} \neq a^o$. We now seek to show that

$$a^{**} > a^o.$$

We do this by contradiction.

We can rewrite (51) as

$$(A.52) \qquad W = (1 - \rho)\lambda u + \rho(1 - \lambda)v$$

$$= \lambda[(1 - \rho)u - \rho v] + \rho v.$$

From (A.52) and noting (53), we can see that if $a$ decreases from $a^o$, then: (i) holding $\lambda$ constant, $[(1 - \rho)u - \rho v]$ will decrease since $u$ achieves maximum at $a^o$ and $v$ is independent of $a$; (ii) holding $[(1 - \rho)u - \rho v]$ constant, $\lambda$ will decrease as $\frac{\partial \lambda}{\partial a} > 0$. Since both $\lambda$ and $[(1 - \rho)u - \rho v]$ are positive, $\lambda[(1 - \rho)u - \rho v]$ will decrease if $a$ decreases from $a^o$. Thus, $a$ must increase from $a^o$ in order to achieve social maximum.

Since the structure of the proofs of the remaining parts of Proposition 9 is akin to the above, the proofs are omitted. ∎

## REFERENCES

ACEMOGLU, D., "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality," *Quarterly Journal of Economics* 113 (1998), 1055–89.

BECKER, G., *A Treatise of the Family* (Cambridge, MA: Harvard University Press, 1991).

——, AND H. G. LEWIS, "On the Interaction between the Quantity and Quality of Children," *Journal of Political Economy* 81 (1973), S279–88.

BENABOU, R., "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review* 86 (1996), 584–609.

BLACKORBY, C., W. BOSSERT, AND D. DONALDSON, "Intertemporal Population Ethics: Critical-Level Utilitarian Principles," *Econometrica* 63 (1995), 1303–20.

——, ——, AND ——, "Utilitarianism and the Theory of Justice," in K. Arrow, A. Sen, and K. Suzumura, eds., *Handbook of Social Choice and Welfare* (Amsterdam: Elsevier, 2002), Vol. 1, 543–96.

BLOOM, B., *All Our Children Learning* (New York: McGraw-Hill, 1981).

BOUCEKKINE, R., D. DE LA CROIX, AND O. LICANDRO, "Vintage Human Capital, Demographic Trends, and Endogenous Growth," *Journal of Economic Theory* 104 (2002), 340–75.

BOULDING, K. E., *The Meaning of the Twentieth Century* (London: Allen & Unwin, 1965).

CHAKRABORTY, S., AND M. DAS, "Mortality, Human Capital and Persistent Inequality," *Journal of Economic Growth* 10 (2005), 159–92.

CURRIE, J., "Early Childhood Education Programs," *Journal of Economic Perspectives* 15 (2001), 213–38.

——, AND D. THOMAS, "Does Head Start Make a Difference?" *American Economic Review* 85 (1995), 341–64.

DAHAN, M., AND D. TSIDDON, "Demographic Transition, Income Distribution, and Economic Growth," *Journal of Economic Growth* 3 (1998), 29–52.

DASGUPTA, P., "The Population Problem: Theory and Evidence," *Journal of Economic Literature* 33 (1995), 1879–902.

DE LA CROIX, D., AND M. DOEPKE, "Inequality and Growth: Why Differential Fertility Matters," *American Economic Review* 93 (2003), 1091–113.

——, AND ——, "Public versus Private Education When Differential Fertility Matters," *Journal of Development Economics* 73 (2004), 607–29.

——, AND A. GOSSERIES, "Population Policy through Tradable Procreation Entitlements," CORE Discussion Paper No. 2006/81, October 2006. *International Economic Review*, forthcoming.

——, AND ——, "Procreation, Migration, and Tradable Quotas," in R. Clark, A. Mason, and N. Ogawa, eds., *Population Aging, Intergenerational Transfers and the Macroeconomy* (Cheltenham: Edward Elgar, 2007), 227–49.

——, AND P. MICHEL, *A Theory of Economic Growth: Dynamics and Policy in Overlapping Generations* (New York: Cambridge University Press, 2002).

EHRLICH, I., AND F. LUI, "The Problem of Population and Growth: A Review of the Literature from Malthus to Contemporary Models of Endogenous Population and Endogenous Growth," *Journal of Economic Dynamics and Control* 21 (1997), 205–42.

GALOR, O., AND O. MOAV, "Natural Selection and the Origin of Economic Growth," *Quarterly Journal of Economics* 117 (2002), 1133–92.

——, AND ——, "From Physical to Human Capital Accumulation: Inequality and the Process of Development," *Review of Economic Studies* 71 (2004), 1001–26.

——, AND ——, "Das Human-Kapital: A Theory of the Demise of the Class Structure," *Review of Economic Studies* 73 (2006), 85–117.

——, AND D. TSIDDON, "The Distribution of Human Capital and Economic Growth," *Journal of Economic Growth* 2 (1997), 93–124.

——, AND D. N. WEIL, "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond," *American Economic Review* 90 (2000), 806–28.

——, AND J. ZEIRA, "Income Distribution and Macroeconomics," *Review of Economic Studies* 60 (1993), 35–52.

GARCES, E., D. THOMAS, AND J. CURRIE, "Longer-Term Effects of Head Start," *American Economic Review* 92 (2002), 999–1012.

GIGLIOTTI, G. A., "Total Utility, Overlapping Generations and Optimal Population," *Review of Economic Studies* 50 (1983), 71–86.

HANUSHEK, E., "Measuring Investment in Education," *Journal of Economic Perspectives* 10 (1996), 9–30.

HAZAN, M., AND B. BERDUGO, "Child Labour, Fertility, and Economic Growth," *Economic Journal* 112 (2002), 810–28.

KREMER, M., AND D. L. CHEN, "Income Distribution Dynamics with Endogenous Fertility," *Journal of Economic Growth* 7 (2002), 227–58.

MERLO, A. M., "Whither Political Economy? Theories, Facts and Issues," in R. Blundell, W. Newey, and T. Persson, eds., *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society* (Cambridge: Cambridge University Press, 2006), Vol. 1, 381–421.

MOAV, O., "Cheap Children and the Persistence of Poverty," *Economic Journal* 115 (2005), 88–110.

MOOKHERJEE, D., AND S. NAPEL, "Intergenerational Mobility and Macroeconomic History Dependence," *Journal of Economic Theory* 127 (2007), 49–78.

NERLOVE, M., A. RAZIN, AND E. SADKA, "Bequests and the Size of Population When Population Is Endogenous," *Journal of Political Economy* 92 (1984), 527–31.

——, ——, AND ——, "Population Size: Individual Choice and Social Optima," *Quarterly Journal of Economics* 100 (1985), 321–34.

——, ——, AND ——, "Some Welfare Theoretic Implications of Endogenous Fertility," *International Economic Review* 27 (1986a), 3–31.

——, ——, AND ——, "Endogenous Population with Public Goods and Malthusian Fixed Resources: Efficiency or Market Failure," *International Economic Review* 27 (1986b), 601–9.

——, ——, AND ——, *Household and Economy: Welfare Economics of Endogenous Fertility* (New York: Academic Press, 1988).

OSBERG, L., *Economic Inequality in the United States* (New York: M. E. Sharpe, Inc., 1984).

PHELPS, E., "The Golden Rule of Accumulation: A Fable for Growthmen," *American Economic Review* 51 (1961), 638–42.

RAWLS, J., *A Theory of Justice* (Oxford: Oxford University Press, 1973).

RAZIN, A., AND E. SADKA, *Population Economics* (Cambridge, MA: MIT Press, 1995).

——, AND C. YUEN, "Utilitarian Tradeoff between Population Growth and Income Growth," *Journal of Population Economics* 8 (2004), 81–87.

——, E. SADKA, AND P. SWAGEL, "Tax Burden and Migration: A Political Economy Theory and Evidence," *Journal of Public Economics* 85 (2002), 167–90.

RESTUCCIA, D., AND C. URRUTIA, "Intergenerational Persistence of Earnings: The Role of Early and College Education," *American Economic Review* 94 (2004), 1354–78.

ROMER, D., *Advanced Macroeconomics* (Boston, MA: McGraw-Hill, 2006).

SAHOTA, G. S., "Theories of Personal Income Distribution: A Survey," *Journal of Economic Literature* 16 (1978), 1–55.

SAMUELSON, P. A., "The Optimum Growth Rate for Population," *International Economic Review* 16 (1975), 531–38.

——, "The Optimum Growth Rate for Population: Agreement and Evaluations," *International Economic Review* 17 (1976), 516–25.

SAW, S., *Changes in Fertility Policy of Singapore* (Singapore: Times Academic Press for the Institute of Policy Studies, 1990).

——, *The Population of Singapore* (Singapore: Institute of Southeast Asian Studies, 1999).

SOARES, R. R., "Mortality Reductions, Educational Attainment, and Fertility Choice," *American Economic Review* 95 (2005), 580–601.

VAN DER EYKEN, W., *The Pre-School Years* (Harmondsworth: Penguin, 1977).

VERNON, P. E., *Intelligence: Heredity and Environment* (San Francisco: W. H. Freeman and Company, 1979).

YAP, M. T., "Fertility and Population Policy: The Singapore Experience," *Journal of Population and Social Security* 1 (Suppl.) (2003), 643–58.