

ENDOGENOUS FERTILITY, INEQUALITY AND HUMAN CAPITAL

Dilip Mookherjee
Boston University

Silvia Prina
Case Western Reserve University
and

Debraj Ray
New York University

October 15 2009

ABSTRACT

This paper integrates theories of endogenous fertility and endogenous inequality in order to explore interactions between fertility choice, human capital investments and inequality. We study steady states of an overlapping generations economy with borrowing constraints and two occupations varying in skill and training costs. Parental altruism is dynastic *a la* Barro-Becker, uses a less restrictive specification of fertility preferences than most existing formulations. We show it is consistent with a negative quality-quantity correlation in demand for children. Endogenizing fertility is shown to remedy two problems of endogenous inequality models: steady states are (generically) locally determinate, and can exhibit intergenerational mobility despite lack of any randomness or agent heterogeneity. The model permits detailed analysis of long run effects of human capital, fertility and fiscal policies.

Mookherjee and Ray acknowledge funding from the National Science Foundation under grant nos. SES-0617874 and SES-0617827 respectively.

1. INTRODUCTION

That endogenous fertility has important implications for long-run living standards of economies has been well known since the time of Malthus. More recently this theme has been emphasized by Barro and Becker (1986,1988,1989) who integrate endogenous fertility into a Ramsey model of optimal growth with a single representative agent and perfect capital markets. By their very nature, these papers ignore questions relating to inequality or capital market imperfections. They also give rise to strong and surprising predictions concerning effects of policies on long-run outcomes: temporary policies have no effect in the long-run, and even permanent policies (e.g., concerning costs of child rearing) have no long run impact on fertility and per capita living standards.¹ Problems have also arisen in reconciling the theory with some of the observed evidence in relation to demographic transition, though the recent contribution of Jones and Schoonbroodt (2009) argues that suitable reparametrizations of the model do enable such a reconciliation.

The past two decades have also witnessed the emergence of a separate literature on endogenous inequality and human capital accumulation in the presence of capital market imperfections (Banerjee-Newman (1993), Galor-Zeira (1993), Ljungqvist (1993), Freeman (1996), Aghion and Bolton (1997), Bandopadhyay (1997), Lloyd-Ellis and Bernhardt (2000), Matsuyama (2000, 2003), Ghatak and Jiang (2002) and Mookherjee and Ray (2002, 2003)). The main focus of this literature is dependence of long run living standards on historical inequality. The models are typically characterized by a continuum of steady states varying in the ‘level of development’ (i.e., higher per capita income steady states have higher per capita skill and lower inequality in wages across skilled and unskilled occupations). The continuum of steady states generates the implication that even small, temporary policies have permanent impact on the level of development.²

The endogenous inequality literature, however, abstracts from endogenous fertility: fertility is assumed constant and unchanging.³ Another problem with this literature is that steady states exhibit no intergenerational mobility, unless perturbed with some randomness in income or abilities.⁴ Moreover, the continuum feature of steady states makes it very difficult to analyze the long-run impact of changes in key parameters or policies, without studying transitional dynamics of these models (which are often analytically intractable, necessitating numerical computations which limit intuitive understanding). This severely limits their usefulness for either descriptive or normative purposes.⁵ For

¹However, Becker, Murphy and Tamura (1990) extend the Barro-Becker approach to incorporate increasing returns to human capital in a representative agent economy, where history dependence may arise in the form of multiple steady states.

²This pertains to models with indivisibilities in entry costs across different occupations. If there are a continuum of occupations with continuously varying entry costs, the steady state is typically unique (Mookherjee and Ray (2003)). But the continuum of steady states obtains with finitely many occupations, irrespective of how many there are.

³Exceptions are noted in subsequent paragraphs.

⁴See Banerjee-Newman (1993) or Mookherjee-Napel (2007) for the latter.

⁵See, for example, the analysis of redistributive fiscal policies in a two-occupation setting in Mookherjee-Ray (2008), which necessitates selection of a particular steady state (out of a continuum) to carry out comparative statics. In that paper, the selection is based on limit points of non-steady-state

instance one is not able to derive predictions concerning the impact on long run per capita income, inequality and mobility of various policies pertaining to child-care, child labor, family planning subsidies or gender discrimination in the labor market, apart from redistributive income taxes.

The purpose of this paper is to integrate endogenous fertility into a model of endogenous inequality with indivisibilities in occupation choice. One would expect interesting interactions between fertility, inequality and human capital in a context with capital market imperfections. Given borrowing constraints, children of unskilled parents will tend to be less skilled than those of skilled parents. Fertility differences between skilled and unskilled parents would then imply a decline in per capita skill in the economy, which is likely to cause an increase in wage inequality between the skilled and unskilled occupations and a reduction in per capita income in the next generation. This point has been emphasized in interesting papers by Dahan and Tsiddon (1998), Kremer and Chen (1999, 2002), de la Croix and Doepke (2003) and Doepke (2004). Most of these papers (with the exception of Kremer and Chen) focus on transitional dynamics rather than properties of steady states.

As elaborated further in the next Section, existing models make very restrictive assumptions on preferences to fit the predictions of the model to some key empirical findings concerning fertility behavior, such as the tendency for families with high wages to have fewer and better-educated children. First, most papers assume that parents treat all their children symmetrically with respect to education: the possibility of educating some children but not others is not allowed. Second, and more important, they make strong assumptions on parameters of the utility function to ensure that substitution effects arising from a rise in wages which reduce fertility, outweigh the accompanying income effects which tend to raise fertility. In particular, in the iso-elastic family of utility functions, they need to assume a greater intertemporal substitution elasticity than is associated with log utility. As we show, if the utility function exhibits at least or greater curvature than the log function, the income effects always outweigh the substitution effects, generating a Malthusian pattern for fertility (i.e., an increase in parental wages raises desired quantity of children, for a given decision concerning their quality). This raises the problem of reconciling the Barro-Becker model with frequently observed facts concerning demographic transition, i.e., the tendency for higher wages to be associated with lower fertility.

Our specification of preferences is considerably more general. We allow parents the option of educating some children and not others, and show that such asymmetric treatment is not optimal quite generally. Hence symmetric treatment is explained rather than assumed.

dynamics, which is somewhat arbitrary and requires an analysis of these dynamics. Minor extensions of these models (e.g., to three occupations) render these dynamics very difficult to analyze. Mookherjee-Ray (2008) therefore cross-validate the results of their particular selection procedure in the two occupation setting against the continuum of occupations setting, where the steady state is unique and therefore permits comparative statics.

With regard to the second problem above, we show that the same results concerning steady states obtain not only when utility exhibits negligible curvature and substitution effects always outweigh income effects, *but also* when the utility function exhibits sufficient curvature and the Malthusian property holds. In both cases, families with a higher level of skill have fewer (or the same number of) children in steady state. Even if the Malthusian property operates in the opposite direction, it is overwhelmed by two countervailing factors that generate a negative correlation between desired quantity and quality of children. One is that higher wages tends to raise desired quality, which is associated quite generally with a drop in desired quantity. There is a threshold wage at which families switch discontinuously from a strategy of high fertility-low quality to low fertility-high quality. The other is a general equilibrium effect, which limits the gap in wages between different occupations around the threshold wage. This limits the strength of the Malthusian effect relative to the drop in fertility at the threshold wage. The model can thus be reconciled with the key facts concerning demographic transition, despite the presence of Malthusian effects (for which there is considerable empirical evidence, e.g., Becker (1960), discussed in greater detail in Section 2.)⁶

With regard to steady state properties, we find that incorporating endogenous fertility remedies the two main shortcomings of the endogenous inequality literature: (a) Steady states are (generically) locally determinate and finite in number, thus permitting local comparative statics with respect to parameters or policies. (b) Steady states can exhibit intergenerational mobility, even without any randomness or inter-agent heterogeneity. These results apply for a wide class of specifications of preferences and technology. Moreover, with iso-elastic utility functions ($u(c) = \frac{c^{1-\rho}}{1-\rho}$), if either ρ is positive but small (so that substitution effects always outweigh income effects), or if it is larger than 1 (so that the opposite is the case and the Malthusian property holds), every steady state involves either zero or upward mobility.

In steady states with upward mobility, a fraction of unskilled parents decide to have fewer children and invest in skill for their children. Other unskilled parents have a large number of children and do not invest in their skill; all skilled parents have fewer children and invest in their skill. Upwardly mobile families have the fewest children. Hence there is a negative correlation between fertility and wages at the household level both in the cross-section as well as over time. Note that this applies irrespective of how strong the Malthusian property is, i.e., with arbitrarily large elasticity of marginal utility of consumption and associated income effects.

The main reason for emergence of upward mobility is the following. Skilled families tend to have fewer and better skilled children. In the absence of mobility, the skill ratio in the economy drifts downward for purely demographic reasons: unskilled agents reproduce at higher rates. This can be arrested either by an increasing wage gap between skilled and unskilled large enough to erase the fertility differences (owing to the operation of the Malthusian effect), or the value of education rises enough to make unskilled parents indifferent between educating and not educating their children. In the

⁶Since completing the first draft of this paper, we became aware of the interesting work of Jones and Schoonbroodt (2009) which makes a similar point, though in the context of a representative agent model with perfect capital markets.

former case a steady state without mobility arises; in the latter case upward mobility appears.

Either kind of steady state is defined by an equation — of fertility across skilled and unskilled families when there is no mobility, and of equal utility of unskilled families between educating and not educating their children. This equation pins down the steady state skill ratio locally. Hence steady states are (generically) locally determinate and finite in number.

The scope for steady state multiplicity is thus drastically reduced owing to the endogeneity of fertility, though not entirely eliminated. Multiple steady states can co-exist, for essentially the reasons stressed by the endogenous inequality literature. They are ordered by the ‘level of development’: a steady state with higher per capita income exhibits higher per capita skill, lower wage inequality across occupations, and higher upward mobility. A higher skill ratio can be self-sustaining because it raises unskilled wages. Owing to borrowing constraints this makes it easier for unskilled parents to invest in their childrens’ education.

The local determinacy of steady states implies no scope for small temporary policies or parameter shocks to have a permanent impact. Temporary policies have to be large enough to shift the state variables into the basin of attraction of a distinct steady state.

On the other hand small permanent policies and parameter shocks do have a permanent impact; our model permits a tractable analysis of the resulting long-term developmental effects, using local comparative statics of a given locally determinate steady state. These provide additional explanations for fertility to fall in the course of development, apart from providing a framework for policy analysis. Reductions in education cost, stronger restrictions on child labor and increases in childcare costs raise long term skill, income and reduce wage inequality, while the impact on mobility is ambiguous in general. Falling gender discrimination in the labor market or cheaper contraceptives also have similar effects, if accompanied by offsetting changes in lump-sum transfers that neutralize the resulting income effects. Unconditional transfers funded by progressive income taxes reduce per capita skill, income and raise wage inequality. These adverse effects can be reversed if unconditional transfers are replaced by conditional transfers (e.g., conditional on school enrollment).

The paper is organized as follows. Section 2 describes related literature. Section 3 introduces the model, following which Section 4 describes household behavior, and 5 presents the main results concerning steady state existence and characterization. Section 6 then describes the long-run impact of various policies, and finally Section 7 concludes.

2. RELATED LITERATURE

Kremer and Chen (1999, 2002) are the only papers we are aware of which focuses on long-run steady state properties of models with endogenous fertility-cum-inequality. Their papers use a very special specification of parental preferences: $U = C + \log n$, which are quasi-linear in own consumption C , and n denotes number of children. Parents do not care about quality of children, or about fertility or welfare of further descendants. The absence of income effects implies that rising parental wages exert only substitution effects

on the demand for fertility. Hence fertility is decreasing in parental wages (where part of childcare costs are time spent away from work). In turn this implies unskilled parents have more children than skilled parents. Parental skill is assumed to reduce the costs of education (either owing to in-home teaching, or unmodeled capital market imperfection), whence children of unskilled (resp. skilled) parents tend to be unskilled (resp. skilled). Fertility differentials between the skilled and unskilled then imply lower per capita skill in the next generation. This amplifies the effect of historical inequality on future inequality and per capita income. Multiple steady states can then arise for purely demographic reasons: low per capita skill economies exhibit greater wage inequality and hence greater fertility differentials between occupations, which lead to low per capita skill in the next generation. Examples of such steady state multiplicity are shown for specific examples of technology and preferences. Kremer and Chen show supporting cross-country evidence that fertility differentials between skilled and unskilled are greater in countries with greater income inequality.

Our model allows for a more general specification of parental preferences. We model parental altruism as dynastic *a la* Barro and Becker. Parents care both about quantity and quality of their children and further descendants. Utility is concave rather than linear in own-consumption, as assumed in the endogenous inequality literature. In such a general specification, changes in parental wages create both income and substitution effects. We show that if the curvature (measured by elasticity of marginal utility) of the own-consumption component of utility is at least as pronounced as in a logarithmic function, income effects overwhelm substitution effects. Fertility then exhibits a ‘Malthusian property’: it is rising in parental wages, within a given occupation. However occupational shifts (arising from rising demand for quality) as wages rise are associated with a decline in fertility. The net effect of wage increases on fertility is the total effect of the Malthusian effect and the occupational shift effect, and is necessarily non-monotonic. This greatly complicates the channel by which Kremer and Chen argue fertility differences amplify the effect of historical inequality on future inequality. This raises the question of robustness of their conclusions to reasonable variations on assumptions regarding preferences, which our paper addresses. Our theory explains why the steady state conditions generally imply that the occupational shift effect dominates the Malthusian effect, no matter how strong the Malthusian effects happen to be.

A number of other papers (e.g., Doepke (2004)) also impose strong assumptions on parental preferences that limit the strength of income effects. Doepke uses a Barro-Becker specification of parental altruism, but assumes that utility functions exhibit less curvature than the log function. His specification of childcare costs is also simpler than ours, by assuming absence of a non-time-varying component. Once one allows child bearing or rearing costs to include a component that does not depend on parental wages, even a logarithmic utility function will exhibit income effects that outweigh substitution effects. Again, Malthusian effects are ruled out by imposing unduly strong assumptions on preferences and technology.

Extensive empirical evidence in favor of the Malthusian property was presented by Becker (1960).⁷ For instance, cross-sectional evidence from the 1940 US Census showed fertility rising with incomes beyond certain levels. Among families that planned their size, desired fertility rose monotonically with parental incomes. The actual tendency for fertility to fall with income in the cross-section is explained by Becker by a larger gap between desired and actual fertility for poorer families (owing to a lower tendency for poor families to plan their size with income, and lower awareness of contraceptives). More importantly, within education classes of the head of household, fertility tended to monotonically increase in income of the head in single earner households. Time-series evidence also shows a tendency for procyclicality of birth rates, even after adjusting for fluctuations in the marriage rate.

This raises the question: how can we reconcile evidence of strong Malthusian effects with strong evidence of negative correlation of fertility with mother's wages, as evidenced both in cross-sectional studies or over time (as in the demographic transition)? The answer provided by Becker (1960) was based on factors such as the decline in child mortality, or increases in contraceptive knowledge (associated with wage increases, urbanization and improvements in public health). Our paper provides an answer to this question without bringing in these additional factors.

As explained above, our analysis differs from Dahan and Tsiddon (1998), de la Croix and Doepke (2003) and Doepke (2004) by focusing on the structure of steady states and their dependence on parameters or policies, while they are interested in transitional dynamics out of steady state. Also, they use numerical calibration methods to models with particular utility and production functions, whereas our analysis is purely theoretical and applies to a large class of utility and production functions. Nevertheless, the two approaches are complementary. The model of Doepke (2004), in particular is essentially the same as ours, except that he rules out Malthusian effects by assumption.⁸

3. MODEL

3.1. Technology. As in most endogenous inequality models there are two occupations, skilled and unskilled. Skilled and unskilled labor constitute the two factors of production. The production function is CRS, smooth, strictly increasing, strictly concave, satisfying Inada conditions (implying essentiality of both kinds of labor). So skilled and unskilled wages will adjust in competitive equilibrium to ensure a positive supply of people in both occupations. Let λ denote the fraction of workers who are skilled at any given date. Then the CRS assumption implies that the skilled and unskilled wage w_1, w_0 are functions of λ alone, the former is decreasing and latter increasing. Moreover w_1 tends to zero and ∞ as λ goes to 1 and 0 respectively, with the opposite holding for w_0 .

⁷Schultz (1973) also presents cross-country evidence showing a positive correlation between fertility and father's wages. The correlation with mother's wages is negative, but it does not control for either mother's or child's education.

⁸Moreover, some of our results concerning properties of household behavior and structure of steady states (Lemma 3, Propositions 1, 2 and the classification of steady states preceding Proposition 7) also appear in his paper.

3.2. Households. This is a standard OLG model: there is a continuum of households, with one adult representative at each date $t = 1, 2, \dots$. This representative may be thought of as the mother. In a later section we shall extend it to incorporate fathers.

We adapt the following specification of household utilities from Becker and Barro. Suppose a parent earns a wage w on the labor market and has n children. As in most formulations of endogenous fertility, we shall ignore integer problems and let n be any non-negative real number.⁹ There could be a physical upper bound to fertility, but as we shall see non-negativity of household consumption will itself impose a uniform upper bound to n .

Every child costs $r(w)$ to rear. Part of these are time costs, while others are resource costs. We thus suppose that $r(w) = rw + f$, where r is the fraction of working lifetime spent away from work per child, and f is the resource cost (including child-care, health care, pre- and post-natal care for the mother). In addition every child costs x to educate, where we assume that x is purely denominated in money. In a later section we shall describe a simple unitary model of the household which makes explicit the determinants of the parameters r, f, x on ‘social’ policies or nature of labor force participation of women, gender disparities etc. and examine effects of varying these. Until then we treat these as fixed parameters.

We also abstract from the possibility of child labor for the time being. Section 6 will indicate how the model can be extended to incorporate child labor.

Unlike most formulations we allow parents to educate some children but not others. The fraction of children educated e can form any fraction in $[0, 1]$. Then a parent with wage w who has n children and educates a fraction e of them is left with lifetime consumption equal to

$$w - [rw + f + xe]n,$$

so we write *overall* utility of the parent as

$$U(w - [rw + f + xe]n) + \delta n^\theta [eV_1 + (1 - e)V_0]$$

where U is smooth, strictly increasing, strictly concave and satisfying $u'(0) = \infty$; $\delta \in (0, \frac{1}{r})$ is the discount rate, and V_1, V_0 represent continuation utilities of children. The ‘effective’ discount rate will turn out to be δn^θ . We impose the upper bound $\frac{1}{r}$ on δ to ensure that the value function iterations form contraction mappings.

If utility is always negative (as in the case of an iso-elastic utility function $U(c) = \frac{c^{1-\rho}}{1-\rho}$ with $\rho > 1$), the parameter θ must be negative to ensure that parents prefer to have more children. If utility is always positive (as with iso-elastic utility and $\rho \in (0, 1)$), the parameter θ must be positive to ensure the same property. The parameter θ is thus restricted to be in $(0, 1)$ if utility is positive always and $-\theta \in (0, 1)$ if it is negative always. As Jones and Schoonbroodt (2009) explain, this specification incorporates the

⁹Conceivably similar results can be obtained in a model with integer-valued family size and heterogeneity in parental fertility preferences, where we can interpret the n obtained in the current theory as the average number of children that would arise in the richer model, for the set of parents with given values of other characteristics incorporated in the current formulation. However, this is a question to be addressed by future research.

key requirements that parental utility be increasing on own consumption, in utility of children and in the number of children (but increasing at a diminishing rate with respect to the latter). These conditions cannot be met if utility is somewhere negative and somewhere positive. For this reason we will not consider the log-utility case.

Finally, consumption is constrained to be nonnegative:

$$w(1 - rn) - (f + xe)n \geq 0$$

which imposes an upper bound to fertility:

$$n \leq \frac{1}{r}$$

4. HOUSEHOLD OPTIMIZATION

Take continuation values V_1, V_0 as given. Denote by $n(w, e)$ the optimal choice of n when wealth is w and skill choice is e . Observe that by the FOC with respect to n ,

$$(1) \quad U'(w - [rw + f + xe]n(w, e)) [rw + f + xe] = \theta [n(w, e)]^{\theta-1} [eV_1 + (1 - e)V_0].$$

Given $V_1 > V_0$ and $w > 0$, this FOC defines the function $n(w, e)$ uniquely (note that at $w = 0$ consumption and n must obviously be zero). Moreover, given the Inada conditions,

$$0 < n(w, e) < \frac{1}{r}$$

For now we suppress the dependence of fertility on V_1, V_0 and other parameters. These will be made explicit whenever needed.

Let the utility of the parent be expressed as a function of e alone:

$$W(e) \equiv U(w - [rw + f + xe]n(w, e)) + [n(w, e)]^\theta [eV_1 + (1 - e)V_0].$$

Using the envelope theorem, we have

$$(2) \quad W'(e; w) = [n(w, e)]^\theta (V_1 - V_0) - U'(w - [rw + f + xe]n(w, e)) \cdot xn(w, e)$$

Using (1) we can express this as follows:

$$(3) \quad W'(e; w) = \frac{[n(w, e)]^\theta}{\frac{rw+f}{x} + e} \left[\left\{ \frac{rw+f}{x} + (1 - \theta)e \right\} (V_1 - V_0) - \theta V_0 \right]$$

Note that the sign of this derivative is the same as that of $\left[\left\{ \frac{r(w)+f}{x} + (1 - \theta)e \right\} (V_1 - V_0) - \theta V_0 \right]$, which is increasing in e . There are now three cases to consider.

First, if

$$(4) \quad \frac{rw+f}{x} (V_1 - V_0) - \theta V_0 \geq 0$$

then the slope is non-negative at $e = 0$, and therefore positive for all $e > 0$. Then it is optimal to set $e = 1$.

Second, if

$$(5) \quad \left[\left\{ \frac{rw+f}{x} + (1 - \theta) \right\} (V_1 - V_0) - \theta V_0 \right] \leq 0$$

then the slope is non-positive at $e = 1$ and therefore negative for all $e < 1$: it is optimal to set $e = 0$.

Finally, suppose neither (4) or (5) hold. The slope is negative at $e = 0$ and positive at $e = 1$. In this case, the slope is negative for a range $(0, e^*)$, zero at e^* and positive thereafter. So e^* is then a local minimum, and every interior e is dominated either by $e = 0$ or $e = 1$. The W function is quasi-convex. From this we infer our first result: *parents either educate all children, or do not educate any of them.*

PROPOSITION 1. *It is always optimal to set $e = 0$ or $e = 1$.*

Next, we compare optimal fertility (for any given w) corresponding to $e = 0$ and $e = 1$ when these are both local optima.

LEMMA 1. *Suppose for given w that both $e = 0$ and $e = 1$ are local optima, i.e., neither (4) or (5) hold. Then $n(w, e = 1) < n(w, e = 0)$.*

Proof. The Kuhn-Tucker condition expressing local optimality of $e = 1$ states that

$$(6) \quad U'(w - (rw + f + x)n(w, 1))x \leq [n(w, 1)]^{\theta-1}(V_1 - V_0)$$

If the Lemma is false then $n(w, 1) \geq n(w, 0)$, which implies that

$$(7) \quad [n(w, 1)]^{\theta-1} \leq [n(w, 0)]^{\theta-1}.$$

Then using the concavity of U we have

$$(8) \quad U'(w - (rw + f)n(w, 0)) < U'(w - (rw + f + x)n(w, 1)).$$

Using (6) and (7), the RHS of (8) is less than or equal to $[n(w, 0)]^{\theta-1}(V_1 - V_0)$. Therefore we obtain

$$(9) \quad U'(w - (rw + f)n(w, 0))x < [n(w, 0)]^{\theta-1}(V_1 - V_0)$$

which contradicts the Kuhn-Tucker condition for local optimality of $e = 0$. ■

The preceding result can be viewed as expressing one version of the well-known “quality-quantity” tradeoff. This tradeoff is a fairly general property, as it does not depend on specific functional forms for the utility function.

Next, let us explore how the optimal choice varies across different w . Is it the case that higher w parents are more likely to go for the low quantity-high quality choice?

Multiplying both sides of (1) by $n(w, e)$ we see that

$$n(w, e)^\theta [eV_1 + (1 - e)V_0] = \frac{1}{\theta} U'(w - [rw + f + xe]n(w, e)) [rw + f + xe]n(w, e).$$

Substituting this into the agent’s maximand yields the following equivalent expression for the agent’s value:

$$V(w, e) = U(w - [rw + f + xe]n(w, e)) + \frac{1}{\theta} U'(w - [rw + f + xe]n(w, e)) [rw + f + xe]n(w, e).$$

The following lemma allows us to draw a curious implication from this expression:

LEMMA 2. *The expression*

$$u(w - z) + \frac{u'(w - z)z}{\theta}$$

is strictly increasing in z if utility is positive and $\theta \in (0, 1)$ and is strictly decreasing in z if utility is negative and $-\theta \in (0, 1)$.

Proof. The derivative of the expression above equals

$$\left[\frac{1}{\theta} - 1\right] u'(w - z) - \frac{u''(w - z)z}{\theta},$$

which is strictly positive in z if utility is positive and $\theta \in (0, 1)$ and is strictly negative in z if utility is negative and $-\theta \in (0, 1)$. ■

This yields the following result:

LEMMA 3. *If utility is positive and $\theta \in (0, 1)$, the agent always acts as if she maximizes $[r(w) + xe]n(w, e)$ by choosing e .*

If utility is negative and $-\theta \in (0, 1)$, the agent always acts as if she minimizes $[r(w) + xe]n(w, e)$ by choosing e .

To prove this, simply compare $V(w, e)$ with the expression in Lemma 2.

When utility is positive and $\theta \in (0, 1)$, we therefore obtain the surprising result that *households choice of occupations for their children follows a simple criterion of maximization (not minimization) of total child-related expenditures*. This property plays a key role in the analysis. The effect of higher investment in children in lowering current household consumption is outweighed by higher benefits placed on the future utilities of the children (reflected in the high investment).

When utility is negative and $-\theta \in (0, 1)$, the opposite result is true. The result applies irrespective of the discount rate or any other parameter in the model.

The following proposition is an immediate corollary of Lemma 3:

PROPOSITION 2. *Consider any wage w at which the agent is indifferent between 0 and 1. Then the expenditures on children must be equalized:*

$$(10) \quad [rw + f + x]n(w, 1) = (rw + f)n(w, 0).$$

In particular, $n(w, 1) < n(w, 0)$ at a point of indifference.

Notice, then, that at a switch point, n not only jumps down as we have already proved, but jumps by *exactly* enough so that the total expenditure on children is unchanged.

This proposition, in turn, gives us the desired single-crossing property for e :

PROPOSITION 3. *Suppose that $e = 1$ at some wealth w ; then $e = 1$ for all wealths thereafter.*

Proof. Define

$$\Delta(w) = U(w - [rw + f + x]n(w, 1)) + [n(w, 1)]^\theta V_1 - U(w - (rw + f)n(w, 0)) - [n(w, 0)]^\theta V_0$$

and differentiate with respect to w , evaluating the result at any point of indifference; i.e., at any point such that $\Delta(w) = 0$. By the envelope theorem applied to $n(w, 1)$ and $n(w, 0)$, we have that

$$\Delta'(w) = U'(w - [(rw + f) + x]n(w, 1)) [1 - rn(w, 1)] - U'(w - (rw + f)n(w, 0)) [1 - rn(w, 0)].$$

By Proposition 2, the terms within the two U 's are exactly equal, and the same proposition also informs us that $n(w, 1) < n(w, 0)$. It follows that $\Delta'(w) > 0$, and we are done. ■

Hence we obtain the result that poor parents (upto some threshold w^*) select $e = 0$ and wealthier parents (above threshold w^*) select $e = 1$. The parent with w^* is indifferent between $e = 0$ and $e = 1$. By Lemma 1 the fertility choice will jump down discontinuously at w^* . This is what we refer to as the ‘occupational shift effect’ associated with increased wages, which is always associated with lower fertility and higher educational investment. We now provide conditions for an interior wage level at which the household switches from $e = 0$ to $e = 1$. In doing that, the following restatement of the FOC (1) will be useful. Denote the optimal expenditure on children as a function of their education and parents wage w :

$$E_e(w) \equiv (rw + f + xe)n(w, e)$$

whence $E_e(w)$ can be represented as the unique solution for E in:

$$(11) \quad E^{1-\theta} U'(w - E) = \left[\frac{1}{rw + f + xe} \right]^\theta V_e$$

In what follows keep in mind that V_1, V_0 are treated as fixed parameters by the household, and it is maintained throughout that $V_1 > V_0$. Later we will make the dependence of E on V_1, V_0 explicit. But not now, in order to keep the notation uncluttered.

PROPOSITION 4. (a) For w sufficiently large, $e = 1$ is optimal.

(b) For w sufficiently small, $e = 0$ is optimal if

$$(12) \quad \left(1 + \frac{x}{f}\right)^\theta V_0 > V_1$$

Otherwise if (12) is violated, $e = 1$ is optimal for all w .

(c) If (12) is satisfied, e switches from 0 to 1 at the wage

$$(13) \quad w^* = \frac{1}{r} \left[x \left\{ \left(\frac{V_1}{V_0} \right)^{\frac{1}{\theta}} - 1 \right\}^{-1} - f \right]$$

Proof. Dividing the FOC (11) corresponding to $e = 0$ by the one corresponding to $e = 1$:

$$(14) \quad \frac{E_0^{1-\theta} U'(w - E_0)}{E_1^{1-\theta} U'(w - E_1)} = \left(\frac{r + \frac{f+x}{w}}{r + \frac{f}{w}} \right)^\theta \frac{V_0}{V_1}$$

For (a), in the case in which utility is positive, if the result is false, $E_1 \leq E_0$ for all large w , whence the LHS of (14) is bounded below by 1 for all large w , as the agent always maximizes total expenditure on children if utility is positive. While the limit of the RHS as $w \rightarrow \infty$ is $\frac{V_0}{V_1} < 1$, and we obtain a contradiction.

In the case in which utility is negative, if the result is false, $E_1 \geq E_0$ for all large w , whence the LHS of (14) is bounded above by 1 for all large w , as the agent always minimizes total expenditure on children if utility is negative. While the limit of the RHS as $w \rightarrow \infty$ is $\frac{V_0}{V_1} > 1$, and we obtain a contradiction.

For (b), consider the positive utility case first. Note that the RHS of (14) tends to $[1 + \frac{x}{f}]^\theta \frac{V_0}{V_1}$ as $w \rightarrow 0$. Hence (12) implies $E_0 > E_1$ for w sufficiently small, as $EU'(w - E)$ is increasing in E . This implies $e = 0$ is optimal for w sufficiently small. And if (12) is violated, we have $[\frac{rw+f+x}{rw+f}]^\theta \frac{V_0}{V_1} < 1$ for all w , implying $E_0 < E_1$, and hence optimality of $e = 1$ for all w .

In the case in which utility is negative, (12) implies $[\frac{rw+f+x}{rw+f}]^\theta \frac{V_0}{V_1} < 1$, and hence $E_0 < E_1$, for all w sufficiently small. This implies $e = 0$ is optimal for w sufficiently small. And if (12) is violated, we have $[\frac{rw+f+x}{rw+f}]^\theta \frac{V_0}{V_1} > 1$ for all w , implying $E_0 > E_1$ for all w . Then $e = 1$ is optimal for all w . This establishes (b).

Finally (c) follows from the requirement of expenditure equality $E_0 = E_1$ at the switch point wage w^* . ■

Condition (12) has a fairly natural and intuitive interpretation. The value V_0 of uneducated children have to be large enough, adjusting for the lower cost of uneducated children (the factor $(1 + \frac{x}{f})^\theta$), compared with the value V_1 of educated children, in order for sufficiently poor families to not want to educate their children. A steady state must have some families deciding not to educate their children; hence wages (and corresponding continuation values V_0, V_1) must adjust to ensure that this condition is satisfied. Essentially, this condition will impose an upper bound to wage differences between the skilled and unskilled in steady states.

Returning to the household optimization problem *per se*, we now turn to the question how fertility varies with parental wages *within a chosen occupation*. This is answered by the following Proposition.

PROPOSITION 5. (a) *Optimal fertility choice $n(w, e)$ is increasing in w for given e if and only if the elasticity of marginal utility $-\frac{cU''(c)}{U'(c)}$ at the optimum is greater than*

$$(15) \quad \frac{rw - \frac{rn}{1-rn}(f + xe)}{rw + f + xe}$$

(b) *For any e , $n(w, e)$ is increasing in w for w sufficiently close to zero, as long as the elasticity of marginal utility of consumption is bounded away from zero as consumption approaches zero.*

(c) *If the elasticity of marginal utility of consumption is everywhere bigger than one, $n(w, e)$ is increasing in w for all w , for any given e .*

(d) *If ρ the elasticity of marginal utility of consumption is constant and less than one, $n(w, e)$ is either inverse-U shaped or monotonically increasing in w for all w , for any given e . For any given positive wage \underline{w} and given e , there exists $\underline{\rho} \in (0, 1)$ such that $n(w, e)$ is decreasing in w for all $w \geq \underline{w}$ if $\rho < \underline{\rho}$.*

Figures 1, 2 and 3 represent the cases where the elasticity of marginal utility of consumption is everywhere bigger than one, close to zero, and less than one respectively.

Proof. (a) From (1), the FOC with respect to n , note that the right-hand-side $\theta n^{\theta-1}[eV_1 + (1-e)V_0]$ is decreasing in n and independent of w . The left-hand-side $U'((1-rn)w - (f+xe)n)[rw + f + xe]$ is increasing in n , and is decreasing in w if and only if

$$rU'((1-rn)w - (f+xe)n) + (1-rn)(rw + f + xe)U''((1-rn)w - (f+xe)n) < 0$$

a condition which is equivalent to the condition that the elasticity of marginal utility of consumption at $c = (1-rn)w - (f+xe)n$ is bigger than (15).

(b) follows from (a) upon observing that (15) tends to zero as w tends to zero. (c) follows from (a) by observing that (15) is everywhere less than one.

The first part of (d) follows from the following observation: expression (15) with $n = n(w, e)$ is increasing in w for given e if $n(w, e)$ is decreasing in w . From (a) we know that if $n(w, e)$ is decreasing in w at some w , the elasticity of marginal utility of consumption ρ is smaller than expression (15). Combining these two observations it follows that if $n(w, e)$ is decreasing in w at some \tilde{w} it must be decreasing in w at all $w > \tilde{w}$. We already know from (b) that $n(w, e)$ is increasing in w for w sufficiently close to zero. Hence either $n(w, e)$ is everywhere increasing in w , or inverse-U shaped.

Finally the second part of (d) follows from noting that expression (15) is bounded away from 0 for all $w \geq \underline{w} > 0$. Otherwise we can find a sequence $\{w_m, e_m, n_m\}_{m=1,2,\dots}$ such that the corresponding value of (15) is converging to 0. Expressing limits by $**$, note that (15) equals $1 - \frac{1}{1-rn^{**}} \frac{f+xe^{**}}{rw^{**}+f+xe^{**}}$. Hence $f+xe^{**}$ must equal $(1-rn^{**})(rw^{**}+f+xe^{**})$. This implies the limiting value of consumption must be $c^{**} = w^{**}(1-rn^{**}) - n^{**}(f+xe^{**})$, which in turn equals $(1-rn^{**})c^{**}$, upon using the fact expressed in the previous sentence. This is possible only if $c^{**} = 0$. This is not possible given that w^{**} is bounded away from 0 and the fact that marginal utility of consumption at 0 is infinite. ■

This Proposition says that optimal fertility must exhibit non-monotonicity with respect to the wage. Part (b) says that for wages sufficiently close to zero, the income effect must overwhelm the substitution effect, as long as the utility function has some curvature throughout. This is intuitive: when the household is very poor, income effects associated with wage increases are strong, and substitution effects are weak (since there are relatively low costs of missing work). On the other hand, from the previous Propositions, we know that the ‘occupational shift’ effect will arise at some interior wage w^* (provided (12) holds, where fertility must locally decline in w).

Part (c) says that if the utility function displays at least as much curvature as the log function, the Malthusian property holds. This condition suffices to ensure that income effects always overwhelm substitution effects.

Part (d) considers the case where utility has constant elasticity and exhibits less curvature than the log function. The preceding arguments imply that a curvature close enough to that of the log function will also give rise to a Malthusian property (except possibly at very large wages). The first part of (d) says it either exhibits this property throughout, or is inverse-U shaped. In the latter case, once the wage crosses a certain threshold,

substitution effects dominate income effects and fertility declines with respect to wage thereafter. The second part of (d) states that this is the case for wages above any positive threshold, if the elasticity of substitution is low enough.

It follows therefore that relying on the absence of the Malthusian property for most of the wage range (i.e., above some low threshold) will require one to assume that the elasticity of marginal utility of consumption is close enough to zero. It seems awkward to base the theory of endogenous fertility on extraordinarily low curvatures of the utility function, as Jones and Schoonbroodt (2009) have also noted in a context of perfect capital markets. They point out that such high elasticities of intertemporal substitution are at odds with what is assumed by most economists in growth and business cycle theory.

In the rest of the paper, we impose no upper bound on the curvature of the utility consumption, thus permitting a pervasive Malthusian property rather than just the vicinity of zero wages. We proceed now to the following Section which considers equilibria and steady states.

5. STEADY STATES

DEFINITION 1. *A **dynamic competitive equilibrium** for the economy starting with skill ratio λ_0 in generation 0 is a sequence of skill ratios $\lambda_1, \lambda_2, \dots$ and associated wages and values $w_{1t}, V_{1t}, w_{0t}, V_{0t}$ of skilled and unskilled workers respectively, satisfying the following properties at every $t = 1, 2, \dots$:*

(a) $w_{1t} = w_1(\lambda_t), w_{0t} = w_0(\lambda_t)$

(b) A parent with skill $s \in \{0, 1\}$ in generation t selects e_t^s, n_t^s to

$$(16) \quad \max_{e,n} [U(w_{st} - (rw_{st} + f + xe)n) + \delta n^\theta \{eV_{1,t+1} + (1-e)V_{0,t+1}\}]$$

and the maximized value of (16) equals V_{st} .

(c) There exists $\eta_{st} \in [0, 1]$ for each skill $s \in \{0, 1\}$ (denoting the fraction of skill s households that decide to educate their children in generation t) with the following property. $\eta_{st} \in (0, 1)$ only if the corresponding households are indifferent between choosing $e = 0$ and $e = 1$ in the optimization problem (16). It equals 1 if they strictly prefer $e = 1$ and 0 if they strictly prefer $e = 0$. Then

$$(17) \quad \lambda_{t+1} = \frac{\lambda_t \eta_{1t} n(w_{1t}, 1) + (1 - \lambda_t) \eta_{0t} n(w_{0t}, 1)}{\lambda_t [\eta_{1t} n(w_{1t}, 1) + (1 - \eta_{1t}) n(w_{1t}, 0)] + (1 - \lambda_t) [\eta_{0t} n(w_{0t}, 1) + (1 - \eta_{0t}) n(w_{0t}, 0)]}$$

DEFINITION 2. *A **steady state** is a dynamic competitive equilibrium with a stationary skill ratio: $\lambda_t = \lambda$ for all t .*

In a steady state, wages and values are also stationary. Note the following features that every steady state must satisfy.

LEMMA 4. *Every steady state has the following properties.*

(a) $w_1 > w_0$ if and only if $V_1 > V_0$.

(b) $V_1 > V_0$, $w_1 > w_0$ and $w_1 - (rw_1 + f + x)n(w_1, 1) > w_0 - (rw_0 + f)n(w_0, 0)$.

(c) $w_1 \geq w^* \geq w_0$, i.e., it is optimal for skilled parents to educate their children, and for unskilled parents to not educate their children.

(d)

$$(18) \quad V_e = \frac{U(w_e - (rw_e + f + xe)n(w_e, e))}{1 - \delta\{n(w_e, e)\}^\theta}$$

These follow from well-known arguments in the endogenous inequality literature. (a) follows from noting the comparison between the optimization problems faced by skilled and unskilled parents which differ only in the parent's current wage. (b) follows from the need to provide some skilled parents incentives to educate their children, combined with their inability to borrow to finance education. Children who receive education must earn higher utilities, which require them to earn higher wages and consume more. This in turn implies that the marginal sacrifice entailed in educating children must be strictly lower for skilled parents. So skilled parents have stronger incentives to invest in education. Part (c) then follows: if skilled parents do not invest, neither will unskilled parents, and there will be no skilled agents in the next generation. Given the essentiality of skill in the production process, a competitive equilibrium cannot have a zero skill ratio. Moreover, it must be optimal for unskilled parents to not invest, otherwise there will be no unskilled individuals in the economy at the next date (whence (b) is violated). Part (d) follows directly from (c).

Note that by virtue of (c):

$$(19) \quad V_e = \max_n [U(w_e - (rw_e + f + xe)n) + \delta n^\theta V_e]$$

Remembering that the choice set in (18) can be restricted to those satisfying $1 - rn > 0$, it follows upon applying the Envelope Theorem that V_e is a strictly increasing function of w_e .

In the analysis of steady states it is important to remember that fertility choices depend not only on parental wages and human capital choices made for children, but also on the continuation values V_1, V_0 . So the RHS of (19) is also a function of V_1, V_0 . Nevertheless, it represents a contraction mapping from continuation values to current values, which thereby has a unique solution. It is easily checked (using property (c) above) that at this solution V_e is a function of w_e alone (i.e., given parameters (δ, r, f, x) , for each e). Since w_e is a function in turn of λ , it follows that V_e can be written as a function of λ alone, with V_1 decreasing in λ and V_0 increasing in λ . By virtue of property (a), we can restrict attention to skill ratios in the interval $(0, \tilde{\lambda})$, where $\tilde{\lambda}$ is defined by the property that $w_1(\tilde{\lambda}) = w_0(\tilde{\lambda})$.

We shall now define some functions which will play a key role in the analysis of steady states.

Note that $n(w, e, V_e)$ has the property that desired fertility depends on V_e alone, and is independent of $V_{e'}, e' \neq e$. Moreover, $n(w, e, V_e)$ is strictly increasing in V_e . Hence we can define $N_e(\lambda) \equiv n(w_e(\lambda), e, V_e(\lambda))$. Given that n is increasing in its first argument, it follows that $N_1(\lambda)$ is strictly decreasing in λ while $N_0(\lambda)$ is strictly increasing. These

functions represent number of children of skilled and unskilled households that decide to invest and not invest in their children's education respectively in steady state.

Next, define

$$E(w, e, \lambda) \equiv [rw + f + xe]n(w, e, V_e(\lambda))$$

the expenditure on children by a parent with wage w selecting e in a steady state with skill ratio λ . Since fertility is increasing in w and continuation value V_e , it follows that this function is increasing in its first argument. For $e = 1$ it is decreasing in λ , while for $e = 0$ it is increasing in λ . Hence the 'switch point' wage w^* defined by the unique solution to

$$E(w^*, 1, \lambda) = E(w^*, 0, \lambda)$$

is a function of λ alone, and is increasing in λ .

We are now in a position to classify steady states into the following classes:

(a) *Steady State Without Mobility (SSWM)* is a steady state (SS) with the property that every skilled household invests in education and every unskilled household does not. In other words, each skilled parent has $N_1(\lambda)$ children all of whom are skilled, and each unskilled parent has $N_0(\lambda)$ children all of whom are unskilled. The requirement of steady state then implies that

$$(20) \quad N_1(\lambda) = N_0(\lambda)$$

as can be verified from the stationary version of (17) with $\eta_1 = 1, \eta_0 = 0$. Conversely, condition (20) in combination with

$$(21) \quad w_0(\lambda) \leq w^*(\lambda) \leq w_1(\lambda)$$

implies $\lambda < \tilde{\lambda}$ is a SSWM, so these two conditions jointly characterize SSWM.

(b) *Steady State with Upward Mobility (SSUM)* is a SS where a positive fraction of unskilled parents invest, and the remaining do not. This requires unskilled parents to be indifferent between investing and not, i.e.,

$$(22) \quad w^*(\lambda) = w_0(\lambda)$$

In such a steady state $w_1(\lambda) > w_0(\lambda) = w^*(\lambda)$ by virtue of property (a), implying that skilled parents all invest. The steady state condition $\lambda_{t+1} = \lambda_t = \lambda$ in (17) combined with $\eta_1 = 1, \mu \equiv \eta_0 \in (0, 1)$ reduces to

$$\lambda = \frac{\lambda N_1 + (1 - \lambda)\mu n}{\lambda N_1 + (1 - \lambda)[(1 - \mu)N_0 + \mu n]}$$

where $N_1 \equiv N_1(\lambda), N_0 \equiv N_0(\lambda), n \equiv n(w_0(\lambda), 1, V_1(\lambda))$. This enables us to solve for the extent of upward mobility:

$$(23) \quad \mu = \frac{N_0 - N_1}{N_0 + \frac{1-\lambda}{\lambda}n}$$

from which it is evident that such a SS must satisfy the condition that

$$(24) \quad N_0(\lambda) > N_1(\lambda)$$

Conversely, conditions (22) and (24) are sufficient to ensure that $\lambda \leq \tilde{\lambda}$ is a SSUM.

(c) *Steady State with Downward Mobility (SSDM)* is a SS where a positive fraction of skilled parents do not invest. Since some of them must invest, skilled parents are indifferent between investing and not. So it must be the case that

$$(25) \quad w_1(\lambda) = w^*(\lambda)$$

In addition an argument analogous to that for SSUM shows it is necessary that

$$(26) \quad N_1(\lambda) > N_0(\lambda)$$

while (26) along with (25) and $\lambda < \tilde{\lambda}$ is sufficient.

Finally, it is evident there cannot be a SS with both downward and upward mobility, since this would require that both skilled and unskilled households are indifferent between investing and not (which would require their respective wages to be the same, equal to $w^*(\lambda)$, contradicting property (b) above). We summarize the preceding discussion:

PROPOSITION 6. *The skill ratio λ is a:*

(a) *SSWM if and only if $\lambda \in (0, \tilde{\lambda})$, and satisfies (20, 21).*

(b) *SSUM if and only if $\lambda \in (0, \tilde{\lambda})$, and satisfies (22, 24).*

(c) *SSDM if and only if $\lambda \in (0, \tilde{\lambda})$, and satisfies (25, 26).*

These are the only types of SS's possible.

We are now in a position to state our main result concerning existence and characterization of SS. In order to do that we need the following definition. Define $\bar{\lambda} < \tilde{\lambda}$ by the condition that

$$(27) \quad w_1(\bar{\lambda}) = w^*(\bar{\lambda})$$

Such a skill ratio is well-defined and unique, since: (i) $w_1(\lambda)$ is strictly decreasing on $(0, \tilde{\lambda})$, (ii) $w^*(\lambda)$ is strictly increasing over this interval, (iii) $w_1(\lambda)$ tends to ∞ as λ tends to 0, and (iv) $w^*(\lambda)$ tends to ∞ as λ tends to $\tilde{\lambda}$ (since in this case the incentive to invest tends to vanish, implying $e = 1$ is optimal for a vanishing set of wage rates).

PROPOSITION 7. (a) *A steady state always exists. Steady state skill ratios are (generically) locally unique and finite in number.*

(b) *If*

$$(28) \quad N_1(\bar{\lambda}) \geq N_0(\bar{\lambda})$$

there is a unique SS at $\lambda = \bar{\lambda}$, which has downward mobility iff (28) holds as a strict inequality.

(c) *If (28) does not hold, there exists a unique $\hat{\lambda} \in (0, \bar{\lambda})$ where*

$$(29) \quad N_1(\hat{\lambda}) = N_0(\hat{\lambda}).$$

If in addition:

$$(30) \quad w_0(\hat{\lambda}) \leq w^*(\hat{\lambda})$$

then $\hat{\lambda}$ constitutes an SSWM. In this case there is no other SSWM or any SSDM. A SSUM at $\lambda \in (\hat{\lambda}, \bar{\lambda})$ exists iff

$$(31) \quad w_0(\lambda) = w^*(\lambda).$$

(d) If neither (28) nor (30) hold, there is no SSWM or any SSDM. There is at least one SSUM at $\lambda \in (\hat{\lambda}, \bar{\lambda})$ satisfying condition (31).

Figures 4, 5 and 6 respectively depict the steady states in cases (b), (c) and (d) respectively.

Proof. Consider first case (b), where (28) holds. Then $\bar{\lambda}$ is a SS where skilled parents are indifferent between investing and not, while unskilled parents strictly prefer not to invest. If (28) holds as an equality, $\bar{\lambda}$ constitutes an SSWM, otherwise it is an SSDM. No other skill ratio can constitute a SS. To see this, note that $\lambda > \bar{\lambda}$ implies $w_1(\lambda) < w^*(\lambda)$, i.e., skilled parents strictly prefer not to invest. On the other hand, at any $\lambda < \bar{\lambda}$ we have $N_1(\lambda) > N_0(\lambda)$ by virtue of (28) and the property that $N_1(\lambda)$ is strictly decreasing while $N_0(\lambda)$ is strictly increasing over $(0, \bar{\lambda})$. So we cannot have a SSWM or SSUM at any such λ . Nor can it be a SSDM, since $\lambda < \bar{\lambda}$ implies skilled parents strictly prefer to invest.

Now consider (c). Note that $w_0(\lambda)$ and hence $N_0(\lambda)$ tends to 0 as λ tends to 0. So $N_1(\lambda) > N_0(\lambda)$ for λ sufficiently small. So if (28) does not hold, there exists unique $\hat{\lambda} \in (0, \bar{\lambda})$ where (29) holds. If (30) also holds, it is clear this is an SSWM. In this case, there cannot be any other SSWM since N_1 is strictly decreasing while N_0 is strictly increasing over $(0, \bar{\lambda})$. There cannot be a SSDM since that requires $w_1(\lambda) = w^*(\lambda)$, i.e., $\lambda = \bar{\lambda}$. Since (28) does not hold, $\bar{\lambda}$ does not constitute a SSDM. A SSUM requires $N_0(\lambda) > N_1(\lambda)$ and $w_1(\lambda) \geq w^*(\lambda)$, so it must involve $\lambda \in (\hat{\lambda}, \bar{\lambda})$ and must also satisfy (31).

Turn to (d) now. If (28) does not hold, the same argument as above establishes no SSDM can exist. Moreover $\hat{\lambda}$ exists as defined above. With (30) not holding, unskilled parents strictly prefer to invest at $\hat{\lambda}$, so there cannot be a SSWM. Now by hypothesis $w_0(\hat{\lambda}) > w^*(\hat{\lambda})$, while $w_0(\bar{\lambda}) < w^*(\bar{\lambda}) = w_1(\bar{\lambda})$. Since these functions are continuous, there exists at least one $\lambda \in (\hat{\lambda}, \bar{\lambda})$ with $w_0(\lambda) = w^*(\lambda)$. At any such λ we have $N_0(\lambda) > N_1(\lambda)$, so this is a SSUM. That there cannot exist any other type of SS in this case can be checked by the reader using variations on the above arguments.

Finally (a) follows upon combining (b),(c),(d). There is at most one $\lambda = \hat{\lambda}$ which constitutes a SSWM. And at most one $\lambda = \bar{\lambda}$ which constitutes a SSDM. As for SSUM, it must satisfy the condition $w_0(\lambda) = w^*(\lambda)$. Raising x the cost of education raises the w^* function everywhere on $(0, \bar{\lambda})$ ¹⁰ while leaving the w_0 function unchanged. Standard transversality arguments (combined with smoothness of the concerned functions) then imply that SSUM skill ratios are (generically) locally unique and finite in number. ■

¹⁰We show this in the next Section.

The steady state characterization generates a number of interesting implications:

- (a) An SSWM and SSUM may co-exist. Then the SSUM must involve a higher skill ratio. It corresponds to a higher ‘level of development’ as it involves higher per capita income, higher per capita skill, lower skill premium in wages, and higher mobility.
- (b) If multiple SSUM’s exist, they are also ordered the same way: a SSUM with higher λ also involves higher upward mobility. This can be checked by differentiating the expression (23) with respect to λ .
- (c) An SSDM cannot co-exist with any other type of SS.
- (d) In a SSUM, there is a negative cross-sectional correlation between parental wages and fertility, and a positive cross-sectional correlation between wages and investment in education. All skilled parents educate their children, while only a fraction of unskilled parents do. Of the latter, those that do not educate their children exhibit the highest fertility. The average education of children of unskilled families is lower than that of the skilled parents.
- (e) In a SSWM, there is no correlation between wages and fertility, and a positive correlation between wages and human capital investment. In a SSDM, there is a positive correlation between wages and fertility, and between wages and human capital investment. In SSWM and SSDM a negative correlation between quantity and quality of children is not exhibited.

We see therefore that in order for the theory to rule out a positive fertility-wage correlation, we need to rule out existence of steady states with downward mobility. To address this question, we specialize to the case of CES utility. It turns out we can rule out downward mobility steady states for a large range of elasticities.

PROPOSITION 8. *Suppose utility is isoelastic $u(c) = \frac{c^{1-\rho}}{1-\rho}$, where $\rho > 0, \neq 1$. Suppose either $\rho > 1$ and $\theta < 0$, or $\theta > 0$ and $\rho \in (0, 1 - \theta)$. Then a steady state with downward mobility cannot exist.*

Proof. In a steady state with downward mobility, it must be the case that $N_1 > N_0$. Denote by \hat{N}_0 the number of children that an occupation-1 parent would choose if she were switching the kids to occupation 0. Parents in occupation 1 must be indifferent between continuing their children in occupation 1 and shifting them to occupation 0 (using r_0 to denote $r \cdot w_0 + f$, and $r_1 = r \cdot w_1 + f + x$):

$$u(c_1) + \frac{\delta N_1^\theta u(c_1)}{1 - \delta N_1^\theta} = u(w_1 - r_0 \hat{N}_0) + \frac{\delta \hat{N}_0^\theta u(c_0)}{1 - \delta N_0^\theta} = u(c_1) + \frac{\delta \hat{N}_0^\theta u(c_0)}{1 - \delta N_0^\theta},$$

where the second equality follows from the fact that total expenditure on children must be the same at the switch point. Therefore

$$(32) \quad \frac{\delta N_1^\theta u(c_1)}{1 - \delta N_1^\theta} = \frac{\delta \hat{N}_0^\theta u(c_0)}{1 - \delta N_0^\theta}.$$

Next, use the first-order condition of parents with wages w_0 and w_1 respectively, in choosing fertility N_0 and \hat{N}_0 :

$$(33) \quad u'(c_0).r_0.N_0^{1-\theta} = u'(c_1).r_1.\hat{N}_0^{1-\theta} = \frac{\theta\delta u(c_0)}{1 - \delta N_0^\theta}$$

Using the iso-elastic expression for u , (33) reduces to

$$\frac{N_0^{1-\theta}}{\hat{N}_0^{1-\theta}} = \frac{r_1}{r_0} \left(\frac{c_1}{c_0}\right)^{-\rho},$$

which implies that

$$(34) \quad \left[\frac{c_1}{c_0}\right]^{\rho(1-\rho)} = \frac{\hat{N}_0^{(1-\theta)(1-\rho)}}{N_0^{(1-\theta)(1-\rho)}} \left[\frac{r_1}{r_0}\right]^{1-\rho}.$$

Similarly, (32) yields

$$\frac{\hat{N}_0^\theta(1 - \delta N_1^\theta)}{N_1^\theta(1 - \delta N_0^\theta)} = \frac{u(c_1)}{u(c_0)} = \left(\frac{c_1}{c_0}\right)^{1-\rho},$$

which upon being raised to the power of ρ yields

$$(35) \quad \left[\frac{c_1}{c_0}\right]^{\rho(1-\rho)} = \left[\frac{\hat{N}_0^\theta}{N_1^\theta} \frac{1 - \delta N_1^\theta}{1 - \delta N_0^\theta}\right]^\rho.$$

Combining (34) and (35):

$$(36) \quad \left(\frac{\hat{N}_0}{N_0}\right)^{1-\theta-\rho} = \left[\frac{N_0^\theta(1 - \delta N_1^\theta)}{N_1^\theta(1 - \delta N_0^\theta)}\right]^\rho \left[\frac{r_0}{r_1}\right]^{1-\rho}.$$

Suppose the result is false and therefore $N_1 > N_0$. Consider first the negative utility case where $\rho > 1$, and $\theta < 0$. Then the right hand side is strictly larger than 1, while the left hand side of (36) is smaller than 1. This is a contradiction. Next in the positive utility case we have $\rho \in (0, 1), \theta > 0$. Then the right hand side is smaller than one and the left hand side is bigger than one provided $\rho < 1 - \theta$. ■

This result says in the case of iso-elastic and negative utility (with a lower intertemporal elasticity of substitution compared with the logarithmic utility function), represented by $\rho > 1$ and $\theta < 0$, a steady state with downward mobility cannot exist. The same result obtains if utility is positive and ρ is close to 0. In the latter case it is less surprising since the Malthusian property is less likely to hold. In the former case the Malthusian property holds everywhere. Despite fertility rising in wages (for any given occupational choice for children), in steady state fertility must exhibit a non-positive correlation with wages. Hence the Malthusian property is overwhelmed by the drop in fertility that must accompany an occupational switch from low to high skill.

For these reasons, it makes sense from this point onwards to focus on SSUM's, and view an SSWM as an extreme form of these where the extent of mobility is zero. Before going on to comparative statics properties of an SSUM, it is convenient to define a SSUM λ to

be *locally stable* if the w_0 curve cuts the w^* curve from above at λ , i.e., the w_0 curve has a lower slope at λ than w^* . If the opposite is true, say that it is *locally unstable*. Figure 6 pictures a locally unstable SSUM. At a locally unstable SSUM, a slight increase in the skill ratio results in w_0 going above w^* , whence all households in the economy strictly prefer to invest, and there is no tendency for the skill ratio to return to the SS. For the comparative statics, we therefore confine attention to locally stable SSUM's. It is easy to check that a locally stable SSUM exists whenever an SSUM exists.

6. COMPARATIVE STATICS

6.1. Reduction in Education Cost. Suppose there is a small decrease in x . While a formal argument can be provided along the lines of Proposition 9 below, we provide a more intuitive argument here. It is clear in this case that V_1 increases at any given λ , while V_0 is unchanged. Moreover, from (11) it follows that $E_1(w)$ rises at every $w > 0$, while $E_0(w)$ is unchanged. From these two facts, it follows (using Lemma 3) that $w^*(\lambda)$ falls at any given λ , i.e., there is a switch at some wages from $e = 0$ to $e = 1$. Hence the w^* curve shifts down, i.e., the incentive to educate increases. Since the w_0 curve is fixed by the technology, it follows that $\bar{\lambda}$ rises (to $\bar{\lambda}'$, say). Moreover, the $N_1(\lambda)$ curve rises, as (using the FOC (1)) the fertility curve $n(w, 1; V_1)$ shifts up, and the V_1 curve has also shifted up. On the other hand the $N_0(\lambda)$ curve is unaffected.

If there is an SSWM to start with at $\hat{\lambda}$, it is possible that an SSWM continues to exist (i.e., if the shifted up N_1 curve continues to intersect the N_0 curve at some $\hat{\lambda}'$ before $\bar{\lambda}'$, and $w_0(\hat{\lambda}') \leq w^*(\hat{\lambda}')$). In that case there is an increase in the skill ratio, while mobility continues to be zero. Fertility rises among both skilled and unskilled households (since the equilibrium results at a higher level of $N_1 = N_0$). *In this case there is an increase in the aggregate population size.* Owing to the constancy of returns to scale, the increase in population *per se* has no effect on per capita income, which rises owing to the increase in per capita skill.¹¹ The effect on population size may appear surprising, as it is commonly expected that cheaper education encourages a shift from quantity to quality of children, thus reducing population size. The reason of course is that in an SSWM there is no correlation between wages or education and fertility. The rise in fertility arises from the Malthusian effects for the unskilled, owing to the rise in unskilled wages. It rises for the skilled owing to the fall in education cost (which overcomes the reverse Malthusian effect arising from lower skilled wages). The conventional intuition would be restored in the presence of a negative correlation, i.e., if there were a SSUM to start with, which we shall consider shortly below.

As mentioned previously, another reason for focusing on SSUM's is to examine the effect of education costs on mobility. Note that in the case where there is an SSWM to start with (at a higher x) but not after x has fallen, it is possible that a SSUM may now exist. This happens when the shifted up N_1 curve continues to intersect the N_0 curve at some $\hat{\lambda}'$ before $\bar{\lambda}'$, and $w_0(\hat{\lambda}') > w^*(\hat{\lambda}')$. *In this case, mobility rises* from zero to a positive level. And the skill ratio rises. The effect on fertility of skilled households is ambiguous (their fertility curve shifts up owing to the reduction in x and the shift

¹¹If there were a fixed factor in the production process, some of the growth would be choked off.

up of the V_1 curve, but this could be counteracted by the decline in the skilled wage resulting from the rise in the skill ratio). An additional reason the aggregate fertility effect is ambiguous is that there are now unskilled households that decide to invest in their childrens' education, and they have fewer children than other unskilled or skilled households that experience no mobility. The upwardly mobile households have the lowest fertility of all, and so an increase in the proportion of such households in the economy tends to bring the aggregate fertility rate in the economy down. Counteracting this, of course, is the rise in fertility among unskilled households that do not experience upward mobility, owing to the increase in the unskilled wage.

Now consider the case where there is an SSUM to start with, and also after the fall in education cost. As argued above the w^* curve shifts down while the w_0 curve is unchanged. If the SSUM to start with is locally stable, the result is an increase in the SS skill ratio, just as in all other cases — so we continue to get a rise in skill and income per capita in the economy, while the skill premium declines.

More interesting is the effect on mobility, given by expression (23). *The net effect appears ambiguous*: while N_0 rises owing to the rise in the unskilled wage, the effect on the value of V_1 and hence N_1 in the SS is ambiguous owing to the fall in the skilled wage. The ambiguity persists even in the absence of the pecuniary externality arising through wage changes.¹² So it seems possible that mobility falls at the same time that cross-sectional inequality in wages also falls. This is interesting in light of the empirical findings of Checchi, Ichino and Rustichini (1999), viz. that mobility in Italy appears to be lower than in the US, despite a more extensive public schooling system in Italy and a lower skill premium in wages. The intuitive explanation of this in terms of our model is that mobility arises to counteract the fertility differential between unskilled and skilled households. This differential drops because the fall in education cost raises fertility among skilled households (because they take advantage of educational opportunities), while leaving that in unskilled households unchanged (if the unskilled wage does not rise much).

6.2. Income Redistribution via Taxes and Transfers. We now show that the results obtained in Mookherjee and Ray (2008) with regard to the effect of different tax-transfer schemes continue to apply in this setting with endogenous fertility. In their setting there was some arbitrariness associated with selection of a particular skill ratio from the continuum of SS skill ratios arising because of exogenously constant fertility. With endogenous fertility this arbitrariness disappears, since SS's are generically locally determinate. It is reassuring to see that similar results obtain with endogenous inequality. Moreover, the current setting generates predictions concerning the effects of income taxes or welfare transfers on fertility and mobility.

¹²Suppose the production function exhibits near-infinite elasticity of substitution between skilled and unskilled labor in the neighborhood of the steady state (i.e., the isoquants are locally linear), so the rise in the skill ratio is associated with negligible changes in skilled and unskilled wages. Then N_0 is unaffected, while N_1 and V_1 rise owing to the rise in x . In addition, the skill ratio rises, while $n(w_0, 1, V_1)$ rises owing to the rise in V_1 . Hence the numerator of (23) falls, but the effect on the denominator is ambiguous owing to the fall in $\frac{1-\lambda}{\lambda}$. If the effect of a rise in λ is outweighed by the effect of a rise in $n(w_0, 1, V_1)$, mobility will fall. Otherwise it may rise.

We provide a brief outline here, omitting most details. Consider first an unconditional welfare system paying an income support of σ to unskilled households, which are financed by income taxes levied on skilled households at a constant linear rate τ . In SS, budget balance requires $(1 - \lambda)\sigma = \lambda\tau w_1(\lambda)$, so the size of the income support depends on the skill ratio and tax rate as follows:

$$(37) \quad \sigma = \frac{\lambda}{1 - \lambda} \tau w_1(\lambda)$$

Steady states have similar properties as established in the previous section, except that the values of being unskilled and skilled are now given by

$$(38) \quad V_0 = \max_{n_0} [U(w_0 + \sigma - (rw_0 + f)n_0) + \delta n_0^\theta V_0]$$

$$(39) \quad V_1 = \max_{n_1} [U((1 - \tau)w_1 - (rw_1(1 - \tau) + f + x)n_1) + \delta n_1^\theta V_1]$$

These policies shift the $V_1(\lambda), N_1(\lambda)$ curves down and raise the $V_0(\lambda), N_0(\lambda)$ curves. Investment in education is thereby discouraged: the w^* curve shifts up. If there is an SSWM before and after the change, the SS skill ratio and per capita income will fall. This will raise the skill premium in wages, so the market will undo some of the redistribution. The same is true if we look at the local comparative static effect on a SSUM: the upward shift in the w^* curve will cause the skill ratio to fall. The fertility differential between the unskilled and skilled rises: owing to the force of the income effects unskilled households will have more (unskilled) children and skilled households will have fewer (skilled) children. As a consequence *it is possible that mobility rises*, though the effect is ambiguous in general, as also the effect on total population.¹³

Many of the adverse long-run effects of unconditional income supports to the poor can be avoided with conditional transfers. An example is an education subsidy which is funded by income taxes on the skilled. In this case the V_0 curve is not directly affected, as many unskilled households do not invest in education in the steady state and so cannot take advantage of the subsidies. On the other hand, the V_1 curve shifts up, as skilled households have more options.¹⁴ This encourages investment in education, as a result of which the w^* curve shifts down, and the SS skill ratio rises (hence so does per capita income, while the skill premium declines). Ignoring the effects of subsequent market-induced changes in wages, fertility among the skilled rises and remains unchanged among the unskilled. The fertility differential narrows, so it is possible that mobility falls. The effects are exactly the opposite of an unconditional welfare system.

¹³Total population size tends to rise on account of rising fertility among the unskilled as well as the shift in the economy in the proportion of unskilled households, while it tends to fall owing to declining fertility among the skilled.

¹⁴If the education subsidy is π , lowering the private education cost to $x - \pi$, budget balance requires $\tau w_1(\lambda) = N_1(\lambda)\pi$, i.e., skilled households pay with taxes for the education subsidies they would avail if they chose the same fertility as before. In that case they are as well off as before. But they have the option to have a different number of children, which makes them better off.

6.3. Changes in Childcare Costs. In the following, we use the notation for consumption of a household whose parent and children have skill e and the economy-wide skill ratio is λ : $C_e(\lambda) \equiv w_e(\lambda) - [rw_e(\lambda) + f + xe]N_e(\lambda)$.

PROPOSITION 9. *Consider any skill ratio λ at which skilled households have at least as much consumption and utility, and no more fertility than unskilled households, i.e.*

$$(40) \quad V_1(\lambda) \geq V_0(\lambda), C_1(\lambda) \geq C_0(\lambda), N_1(\lambda) \leq N_0(\lambda).$$

(a) *At any such skill ratio, investment incentives are increasing in f , i.e., $w^*(\lambda)$ is locally decreasing in f .*

(b) *Every (locally stable) SSUM the skill ratio is locally increasing in f .*

Proof. (a) Using the fact that total child expenses are equalized across the options of investing and not at the wage w^* , we obtain from (14):

$$(41) \quad w^*(\lambda) = \frac{1}{r} [x \{ (\frac{V_1(\lambda)}{V_0(\lambda)})^{\frac{1}{\theta}} - 1 \}^{-1} - f]$$

which shows that for a given value of $\frac{V_1}{V_0}$, w^* is decreasing in f . So it suffices to show that $\frac{V_1}{V_0}$ is weakly increasing in f . Recall that

$$(42) \quad V_e(\lambda; f) = \max_{n_e} [U(w_e(\lambda) - (rw_e(\lambda) + f + xe)n_e) + \delta n_e^\theta V_e(\lambda; f)]$$

Applying the Envelope Theorem to this optimization problem we have

$$(43) \quad \frac{\partial V_e(\lambda; f)}{\partial f} = - \frac{U'(C_e(\lambda))N_e(\lambda)}{1 - \delta N_e(\lambda)^\theta}.$$

Therefore (40) implies that

$$V_0(\lambda) \frac{\partial V_1(\lambda; f)}{\partial f} - V_1(\lambda) \frac{\partial V_0(\lambda; f)}{\partial f} = -V_0(\lambda) \frac{U'(C_1(\lambda))N_1(\lambda)}{1 - \delta N_1(\lambda)^\theta} + V_1(\lambda) \frac{U'(C_0(\lambda))N_0(\lambda)}{1 - \delta N_0(\lambda)^\theta} \geq 0.$$

(b) At any SSUM (40) is satisfied. Moreover by (a), the w^* curve shifts down following an increase in f , while the position of the $w_0(\lambda)$ is fixed by the technology. Hence if the SSUM is locally stable the skill ratio must rise. ■

We therefore obtain the interesting result that a rise in component of childcare cost that is unrelated to parental time away from work raises skill formation and per capita income in a SSUM. At the switch point w^* where expenditures associated with investing and not are equalized, it tilts the balance in favor of the investment option. This is for both a direct reason (childcare expenses per child are lower when not investing in education, so a rise in f raises childcare costs by more for non-investors) and an indirect reason (the continuation utility of skilled children falls by less than for unskilled children, owing to (40)). It predicts that societies with a norm where extended family or kinship networks share the burden of child-rearing (so the parents bear a smaller part of the burden) will tend to invest less in education of the children. Social changes that cause a shift from joint to nuclear families thus induce higher education. *Policies of subsidized childcare provided*

by the government harm skill accumulation incentives, and raise inequality between skilled and unskilled wages.

Effects on aggregate fertility are ambiguous. A rise in f tends to lower fertility by shifting both $N_0(\lambda)$ and $N_1(\lambda)$ curves down, as well as by inducing an increase in the proportion of skilled households who have lower fertility. Unskilled wages rise owing to the rise in the skill ratio, so the net effect on the fertility of the unskilled is ambiguous. On the other hand, fertility among the skilled must fall. Since the effects on fertility are ambiguous, so are the effects on mobility.

We have not yet managed to obtain a comparable result with respect to the parameter r , the time component of childcare costs. The direct effect of an increase in r is similar to that of f : both raise w^* , enhancing skill investment incentives. The indirect effect is however more difficult to assess, since it is no longer clear that a rise in r lowers V_1 by less, as this effect is also proportional to the wage rate which is higher for the skilled.

Note also that the result applies only to steady states with upward mobility. It is not clear whether it also applies to steady states without mobility. An increase in f will cause both $N_0(\lambda)$ and $N_1(\lambda)$ curves to shift down. The effect on the steady state skill ratio will depend on which falls by more, and we have not yet been able to sign this. However, for these steady states, one prediction can be made: aggregate fertility must fall, since they are equalized across all households in the economy and the fertility curves for both groups shift down.

6.4. Child Labor. Now suppose children that do not go to school can work and augment the incomes of their households. Suppose that children can work as a substitute for unskilled adult labor, and earn a wage of γw_0 , where $\gamma \in (0, 1)$ is a parameter that reflects differences in work capacity between adults and children, as well as regulations concerning child labor. Stronger restrictions on child work correspond to a reduction in γ . The preceding model can be thought of as corresponding to the case where $\gamma = 0$.

In this case, household consumption corresponding to parental wage w is $C \equiv w - [rw + f + xe - \gamma w_0(1 - e)]n$. This corresponds to our earlier model if we replace f by $f' \equiv f - \gamma w_0$ and x by $x' \equiv x + \gamma w_0$. To ensure $f' > 0$ we must impose the restriction that $\gamma < \frac{f}{w_0(\lambda)}$.

Stronger restrictions on child labor then correspond to a fall in γ , which is analytically equivalent to a rise in child care costs combined with a fall in education cost. Both of these induce a rise in the steady state skill ratio.

There is an additional effect that operates through the effect on wages (stressed in particular by Basu and Van (1998)): a reduction in child labor reduces the supply of unskilled labor in the economy as a whole, which tends to raise unskilled wages. Hence the $w_0(\lambda)$ curve shifts up. This has an additional effect on an SSUM, since it is characterized by intersection of the w^* curve and the w_0 curve. If the SSUM is locally stable to start with, this effect raises the steady state skill ratio even further.

Hence the net effect is to raise the ‘level of development’ in the long run: higher per capita skill and income. Effects on fertility and mobility are ambiguous.

6.5. Family Planning Costs; Gender Discrimination in the Labor Market.

A number of additional complications arise when we extend the model to consider the effect of changes in gender discrimination in the labor market, or changes in family planning subsidies. Consider the following extension of our model to a context where each household has a male and a female parent in each generation.

Assume perfect assortative matching in marriage, so people marry others of the same skill. However, suppose there are some restrictions on pay or work of women, whereby the woman earns a wage of $\alpha \in (0, 1]$ times the male wage at the same skill level, where $1 - \alpha$ is an index of gender discrimination.

In addition assume that both spouses share the same objectives, and make a joint decision concerning children and their education. Their common objective is $U(C) + \delta n^\theta [eV_1 + (1 - e)V_0]$, where C is total household consumption. So for a given decision on fertility (n) and education (e) they will make intra-household decisions to maximize household consumption C .

Each child requires a total parental time equal to h fraction of an adult's working lifetime. If the family decides to have n children, the total time to be devoted to children equals $H \equiv nh$. Clearly H cannot exceed 2, as the family will then have no income to take care of their children.

If $\alpha < 1$, maximizing family income requires minimizing time spent by the father looking after the children. If $H > 1$, the mother has no time to work on the labor market, and the father spends $H - 1$ fraction of his working life at home. In this case total household income equals $w(2 - H)$, locally independent of α . More interesting is the case where $H < 1$, whence the father works fulltime and the mother works $1 - H$ part of her lifetime. Here family income equals $w + \alpha w(1 - H)$, which is rising in α . Lower gender discrimination allows female wages to catch up with male wages and family income to rise (ignoring general equilibrium effects on wage rates).

Pre-school childcare costs g per child, which includes costs of early childhood health care for the child and the mother. Moreover, family planning (contraceptive) costs equal $c(\bar{n} - n)$, where \bar{n} is the natural upper limit to fertility resulting from total lack of control over reproductive decisions. Then total costs of children excluding education expenses equal $gn + c(\bar{n} - n) = (g - c)n + c\bar{n}$. Family consumption then equals (in the case that $nh < 1$): $w + \alpha w(1 - hn) - (g - c)n - xen + c\bar{n}$. We can thus identify the parameter r with αh , and f with $g - c$.

However there is one difference from the previous model: an increase in α both raises r and family income at the same time, assuming there is no net effect on the wage structure.¹⁵ As explained in the discussion following Proposition 9 the effect of an increase in r cannot be signed. This ambiguity is further compounded owing to additional income effects to consider.

Similarly, a family planning subsidy corresponds to a fall in c , which has two effects: it both raises the parameter $f \equiv g - c$, as well as generates an income effect (the fixed cost

¹⁵If α represents a restriction on the hours that women can work, then a decrease in α will result in an increase in female labor force participation, which will affect wages. Then the effect is going to even more complicated.

component $c\bar{n}$). It amounts to a lump-sum subsidy combined with an increase in the incremental cost per child. Hence the effect of a family planning subsidy (fall in c) is not the same as an increase in childcare cost g , as the latter does not involve a corresponding lump-sum income effect. While the direct effect (on w^* for given continuation utilities V_1, V_0 and given skill ratio) of a fall in c on education incentives is in the same direction as that resulting from a rise in g , the indirect effect is different (the lumpsum income effect is valued more by the unskilled, so it is no longer clear that V_1 falls by less than V_0). The size of this income effect is higher, the greater \bar{n} is relative to actual fertility levels. One would expect therefore that in sub-Saharan countries where total fertility rates of women are very high and close to their reproductive capacities, this income effect will be negligible. In those contexts, family planning subsidies will have the same effect as rising childcare costs: they will encourage greater investments in child education. The same will be true if the falling contraceptive costs are compensated with lumpsum taxes that neutralize the income effects.

7. CONCLUDING COMMENTS

We have explored interactions between endogenous fertility and endogenous inequality and argued that this enriches both theories of fertility and inequality, apart from allowing us to study the implications of fertility-related policies on long run inequality and living standards.

While we have managed to reduce the restrictiveness of specification of fertility preferences considerably relative to existing literature, there is scope for further generalizations of the theory (e.g., in the iso-elastic case for all possible levels of marginal utility elasticity, including intermediate levels where it is difficult to sign the relative strengths of income and substitution effects).

Our theory generates a number of empirical predictions, such as the tendency for fertility to drop when parents switch from low to high quality children, the relation between patterns of fertility and mobility, and the long run effects of various policies on human capital, inequality and living standards. We hope this will spur empirical work to test these predictions.

References

- Aghion P. and Bolton P.** (1997). "A Theory of Trickle-Down Growth and Development." *Review of Economic Studies*, 64, 151-172.
- Bandyopadhyay D.** (1997). "Distribution of Human Capital and Economic Growth," Working Paper no. 173, Department of Economics, Auckland Business School.
- Banerjee A. and Newman, A.** (1993). "Occupational Choice and the Process of Development." *Journal of Political Economy* April 1993, 101(2), pp. 274–298.
- Barro R. and G. Becker** (1989), "Fertility Choice in a Model of Economic Growth," *Econometrica*, 57(2), 481-501.
- Basu K. and P. Van** (1998), "The Economics of Child Labor," *American Economic Review*, 88(3), pp. 412-27.
- Becker G.** (1960), "An Economic Analysis of Fertility," in *Demographic and Economic Change in Developed Countries*, Princeton: Princeton University Press.
- Becker G. and R. Barro** (1986), "Altruism and the Economic Theory of Fertility," *Population and Development Review*, 12 Supplement Issue) , 69-76.
- (1988), "A Reformulation of the Economic Theory of Fertility," *Quarterly Journal of Economics*, 103(1), 1–25.
- Becker G. and N. Tomes** (1976), "Child Endowments and the Quantity and Quality of Children," *Journal of Political Economy*, 84(4), Part 2, S143-S162.
- Becker G., K. Murphy and R. Tamura** (1990), "Human Capital, Fertility and Economic Growth," *Journal of Political Economy*, 98(5) Part 2, S12-S37.
- Checchi A., A. Ichino and A. Rustichini** (1999), "More Equal but less Mobile? Education Financing and Intergenerational Mobility in Italy and in the U.S.," *Journal of Public Economics*, 74, 351-393.
- De La Croix D. and M. Doepke** (2003), "Inequality and Growth: Why Differential Fertility Matters," *American Economic Review*, 93(4), 1091–1113.
- Dahan M. and D. Tsiddon** (1998), "Demographic Transition, Income Distribution and Economic Growth," *Journal of Economic Growth*, 3, 29–52.
- Doepke M.** (2004), "Accounting for Fertility Decline During the Transition to Growth," *Journal of Economic Growth*, 9, 347–383.
- Freeman, S.** (1996), "Equilibrium Income Inequality among Identical Agents." *Journal of Political Economy* October 1996, 104(5), pp. 1047–1064.
- Galor, O. and Zeira, J.** (1993), "Income Distribution and Macroeconomics." *Review of Economic Studies* January 1993, 60(1), pp. 35–52.
- Ghatak, M. and N. Jiang** (2002), "A Simple Model of Inequality, Occupational Choice and Development" *Journal of Development Economics* **69**, 205–226.
- Jones L. and A. Schoonbroodt (2009), "Complements versus Substitutes and Trends in Fertility Choice in Dynamic Models," working paper, University of Minnesota.
- Kremer M. and L. Chen** (1999), "Income Distribution Dynamics with Endogenous Fertility," *American Economic Review*, 89(2), 155-160.

————— (2002), “Income Distribution Dynamics with Endogenous Fertility,” *Journal of Economic Growth*, 7, 227-258.

Ljungqvist, L. (1993), “Economic Underdevelopment: The Case of Missing Market for Human Capital,” *Journal of Development Economics*, 40, 219-239.

Lloyd-Ellis H. and Bernhardt D. (2000), “Enterprise, Inequality and Economic Development,” *Review of Economic Studies*, 67(1), 147-168.

Matsuyama, K. (2000), “Endogenous Inequality”, *Review of Economic Studies*, 67, 743-759.

————— (2003), “On the Rise and Fall of Class Societies,” mimeo, Department of Economics, Northwestern University.

Mookherjee D. and Napel S. (2007), “Intergenerational Mobility and Macroeconomic History Dependence,” *Journal of Economic Theory*.

Mookherjee, D. and D. Ray (2002), “Is Equality Stable?,” *American Economic Review* **92** (Papers and Proceedings), 253-259.

————— (2003), “Persistent Inequality,” *Review of Economic Studies*, 70(2), April 2003, 369-394.

Schulz, P. (1973), “A Preliminary Survey of Economic Analysis of Fertility,” *American Economic Review* 63(2) (Papers and Proceedings), 71-78.

Whelpton P.K. and C.V. Kiser (1951), *Social and Psychological Factors Affecting Fertility*, Milbank Memorial Fund, Vol 1-4.

Figures

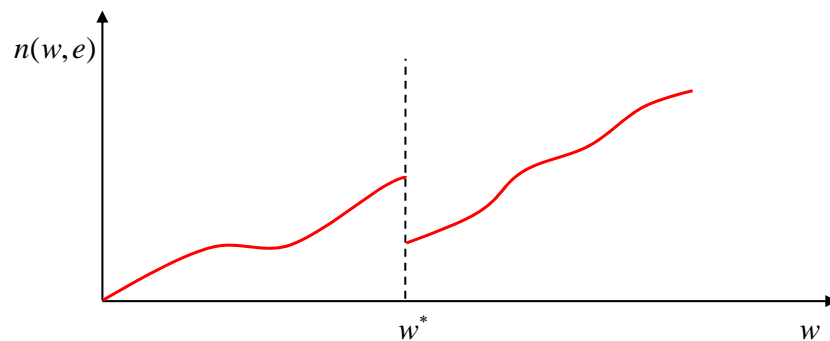


FIGURE 1. Elasticity of marginal utility of consumption bigger than 1.

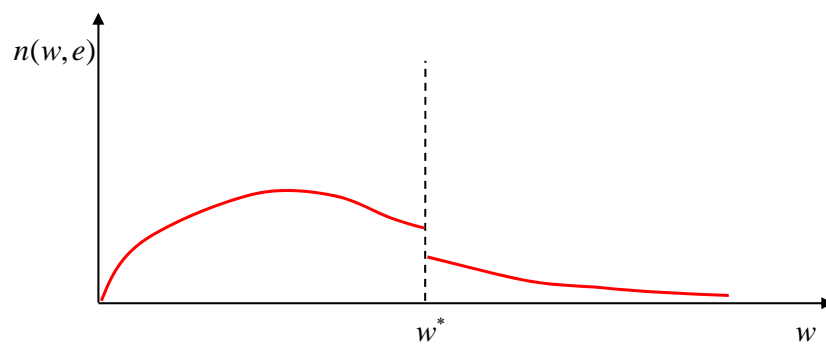


FIGURE 2. Elasticity of marginal utility of consumption close to zero.

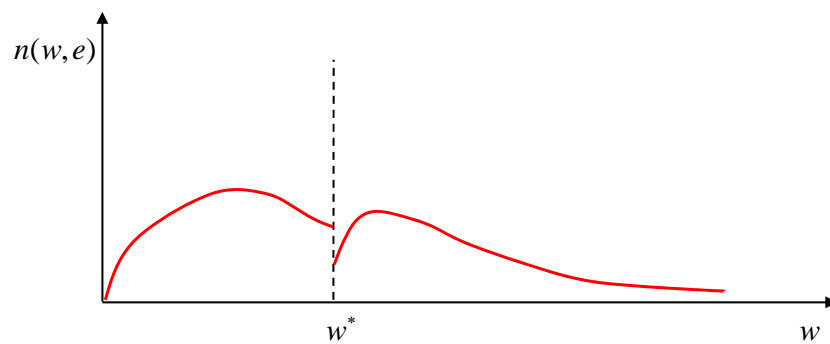


FIGURE 3. Elasticity of marginal utility of consumption intermediate, less than 1.

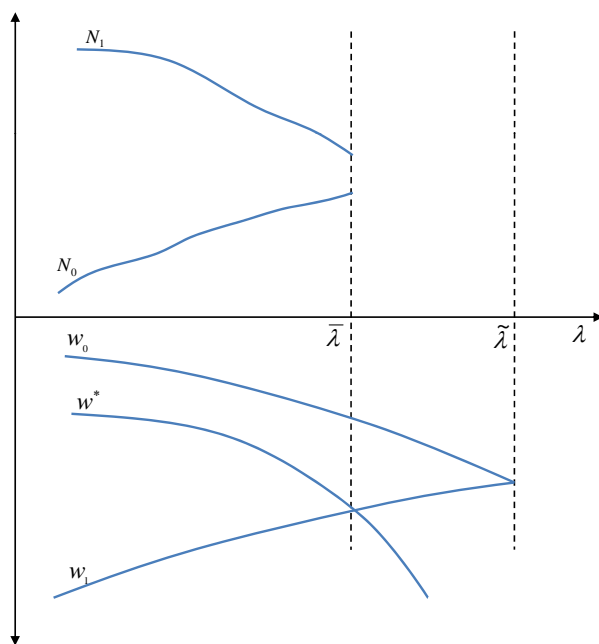


FIGURE 4. Steady state with downward mobility at $\tilde{\lambda}$.

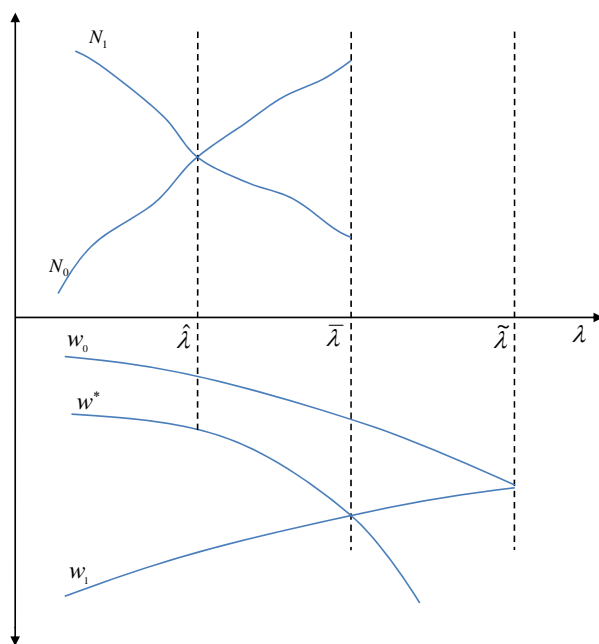


FIGURE 5. Steady state without mobility at $\hat{\lambda}$.

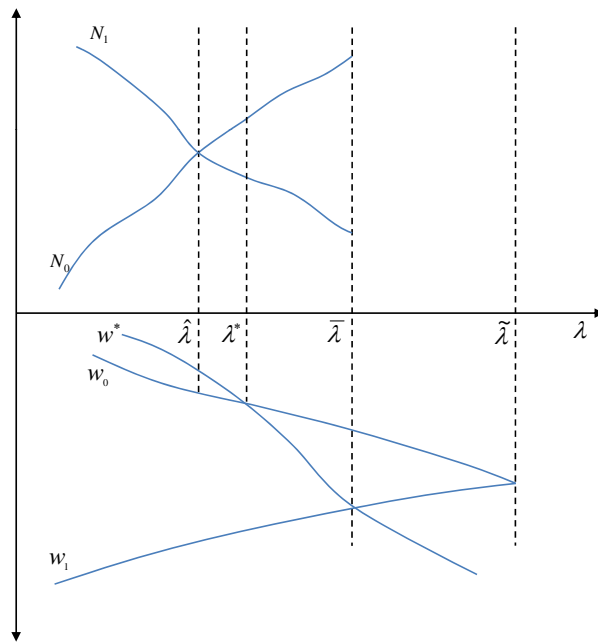


FIGURE 6. Steady state with upward mobility at λ^* .