Journal of Economic Growth, 7, 227–258, 2002

Income Distribution Dynamics with Endogenous Fertility

MICHAEL KREMER

Department of Economics, Harvard University, Littauer Center, Cambridge, MA 02138, USA, Brookings Institution and National Bureau of Economic Research

DANIEL L. CHEN

MIT, Cambridge, MA 02138, USA

Developing countries with highly unequal income distributions, such as Brazil or South Africa, face an uphill battle in reducing inequality. Educated workers in these countries have a much lower birthrate than uneducated workers. Assuming children of educated workers are more likely to become educated, this fertility differential increases the proportion of unskilled workers, reducing their wages, and thus their opportunity cost of having children, creating a vicious cycle. A model incorporating this effect generates multiple steady-state levels of inequality, suggesting that in some circumstances, temporarily increasing access to educational opportunities could permanently reduce inequality. Empirical evidence suggests that the fertility differential between the educated and uneducated is greater in less equal countries, consistent with the model.

Keywords: fertility, inequality, fertility differential

JEL classification: J1, O1

1. Introduction

In developing countries, fertility typically falls with education. For example, in Brazil, women with no education have three times as many children as women with 10 or more years of education (United Nations, 1995). Since children of the uneducated are less likely to become educated themselves, this three-fold difference in fertility creates a major demographic force increasing the proportion of unskilled workers.

One plausible hypothesis for why fertility declines with education is that because educated women command higher market wages, they face higher opportunity costs of time spent rearing children. If substitution effects outweigh income effects, then educated women have fewer children.

This paper examines the implications of combining three assumptions: (i) skilled and unskilled workers are complements in production; (ii) children of the unskilled are more likely to be unskilled; and (iii) higher wages reduce fertility because substitution effects outweigh income effects. A model incorporating these features implies that an initial increase in the proportion of unskilled workers will reduce wages of unskilled workers. Since lower wages decrease the opportunity cost of raising children, unskilled workers will increase their fertility. Under the assumption that children of unskilled workers are more likely to be unskilled themselves, the proportion of unskilled workers in the next generation will therefore increase. An initial increase in the fraction of unskilled workers thus produces a multiplier effect in subsequent generations, suggesting that improving educational opportunities for even small numbers of children of unskilled workers could lead to large changes in the steady-state distribution of skill.

The model also generates multiple steady-state levels of inequality. If the initial proportion of skilled workers is great enough, wage and fertility differentials between skilled and unskilled workers will be small, allowing the economy to converge to a steady state with low inequality. However, if the initial proportion of skilled workers is too low, inequality will be self-reinforcing and the economy may approach a steady state with a low proportion of skilled workers and great inequality between the skilled and unskilled. Increasing the probability that children of unskilled parents will become skilled, for example, by expanding access to educational opportunities, reduces the basin of attraction of the unequal steady state and may even eliminate it. Some countries may face a brief window of opportunity in which small and temporary increases in the probability that children of unskilled parents become skilled can move them into the basin of attraction for the equal steady state. As time passes, however, and the economy approaches the unequal steady state, larger or longer-lasting increases in the probability that children of unskilled parents become skilled will be necessary to move to the more equal steady state.

We find empirical evidence that the fertility differential between educated and uneducated women is greater in countries with more income inequality. Using data on differential fertility by education from the World Fertility Surveys (United Nations, 1987; Jones, 1982) and Demographic and Health Surveys (United Nations, 1995; Mboup and Saha, 1998), we find a statistically significant, positive, and economically sizeable relationship between differential fertility and inequality for most specifications, consistent with the model.

Our model is most applicable to middle-income countries. At very low wages, wage increases may increase the number of surviving children by reducing infant mortality and infertility due to disease and malnutrition. At very high wages, further wage increases are likely to reduce fertility only modestly if fertility asymptotes to a positive level. In fact, the positive relation between inequality and fertility differentials is strongest in middle-income countries.

1.1. Literature Review

Several writers have explored one direction of the feedback mechanism—the impact of differential fertility on the distribution of income or socio-economic status (Lam, 1986; Chu and Koo, 1990; Preston and Campbell, 1993; Mare, 1997). These papers use a Markovian framework in which fertility in each group (for example, skilled and unskilled) and the probability that a child born to parents in one group will transit to another group are both independent of the distribution of the population across groups. Many other papers discuss fertility and income inequality more generally (Repetto, 1978; Galor and Zang, 1997; Nerlove et al., 1984; Galor and Zeira, 1993). None of these papers, however, capture the other direction of causality. As informally discussed by Birdsall (1988), an increase in income inequality, measured by a greater number of unskilled workers relative to skilled workers, is likely to suppress the wages of the unskilled. If fertility depends inversely on wages, this increases the fertility of the unskilled and thus increases fertility differentials, creating a positive feedback. On the other hand, increasing income inequality implies higher wage differentials, which in turn increases the incentive for children of the poor to become educated and transit to a higher income group. This may create a negative feedback effect which counteracts the positive feedback at the steady state.

Dahan and Tsiddon (1998) and Morand (1999) incorporate this positive feedback in papers modeling the demographic transition and Kuznets curve, focusing on transition dynamics. In Dahan and Tsiddon's model, poor dynasties initially do not invest in education and therefore stay poor, whereas rich dynasties invest in education and stay rich. Thus, there is initially no intergenerational mobility between rich and poor. Due to greater fertility among poor, the proportion of poor increases, leading to greater inequality. Once income inequality reaches a certain threshold, wage differentials and thus incentives to obtain education are great enough for some poor to obtain education. The number of uneducated people then falls, which increases their wages and reduces income inequality. At this point, inequality moves to its steady-state level.

Whereas Dahan and Tsiddon (1998) focus on transition dynamics in a model with a single steady state, we show that the positive feedback between fertility differentials and income inequality may lead to multiple steady states and multiplier effects. Dahan and Tsiddon obtain a single steady state because they assume all poor people are identical rather than varying in their cost of education.

Moav (2001) obtains multiple steady states in a model in which educated people have a comparative advantage in producing quality children although he emphasizes dynastic implications.

The remainder of this paper is organized as follows: Section 2 presents the model, solves for the steady states, explores the transition path to the steady state from various initial conditions, and discusses how the dynamic system changes in response to changes in parameters. Section 3 describes the data, discusses the methodology, and presents the empirical results. Section 4 argues that the substitution effect through which wage increases reduce fertility are likely to be important at moderate income levels, and presents evidence that the empirical relationship we document is concentrated in middle-income countries. Section 5 argues that the basic results are robust to various generalizations and extensions of the model and concludes.

2. Model

This section is organized as follows. Section 2.1 describes the model. Section 2.2 solves for the steady states. Section 2.3 discusses stability of the steady states. Section 2.4 explains how the basin of attraction and the wage differentials at the steady state change in response to changes in the parameters.

2.1. Model

Suppose the production technology is

$$Y = A\sqrt{L_t^s\, L_t^u} = A\, L_t^u\, R_t^{1/2}, \tag{1}$$

where L_t^s and L_t^u are the number of skilled and unskilled workers in time t, respectively and R_t, the ratio of skilled to unskilled workers at time t, is L_t^s/L_t^u.[1] A working paper version of this paper considers a general exponent in the production function (Kremer and Chen, 2000).

Assuming competitive factor markets, wages will be $w_t^s = AR_t^{-1/2}/2$ and $w_t^u = AR_t^{1/2}/2$ and therefore the wage differential, $w_t^s/w_t^u = 1/R_t$.

In general, increases in wages create substitution effects, making time more valuable and thus reducing the number of children, and income effects increasing the number of children. Given the negative correlation between wages and fertility, and between education and fertility, both across countries and within countries (except at very low wage levels), we assume substitution effects outweigh income effects.[2] In particular, we assume the fertility function to be

$$n_t^i = 1/(\phi w_t^i) \quad \text{for } i = s, u, \tag{2}$$

where n_t^s and n_t^u are the number of children of skilled and unskilled workers respectively.[3] As discussed in Section 4, the inverse relationship between fertility and wages is likely to be a more acceptable approximation over the moderate wage levels characteristic of middle-income countries than at very low or very high wages.[4]

To complete the model, it is necessary to specify the process governing each individual's education decision. We assume that (i) educational decisions are responsive to the incentives provided by wage premia, and (ii) children of unskilled parents face higher costs of education than children of skilled parents, due to either differences in home environments or lack of access to capital markets. To capture these features, consider a model in which individuals can choose to become skilled or unskilled. To become skilled, each individual i needs to invest c^i units of time in education. Suppose all children of skilled parents, along with a proportion θ of children of unskilled parents need c^L units of time to become skilled. A proportion $1-\theta$ of children of unskilled parents need c^H units of time to become skilled, where $c^L < c^H$.[5] Individual i with time cost c^i will choose to become skilled if the lifetime income from obtaining education is greater than the lifetime income from not obtaining education, that is, when

$$(1 - c^i)w_t^s \geq w_t^u \Leftrightarrow 1 - c^i \geq R_t. \tag{3}$$

Hence, if $R_t = 1 - c^L$, low-cost individuals, with cost c^L, are indifferent between becoming skilled or unskilled and high-cost individuals, with cost c^H, prefer to become unskilled. If $R_t \in (1 - c^H,\ 1 - c^L)$, then low-cost individuals prefer to become skilled and high-cost individuals prefer to become unskilled. If $R_t = (1 - c^H)$, then low-cost individuals prefer to become skilled and high-cost individuals are indifferent.

If $R_t > (1 - c^L)$, no one chooses to become skilled, and if $R_t < (1 - c^H)$, everyone chooses to become skilled; this implies that in equilibrium, $R_t \in [1 - c^H,\ 1 - c^L]$.

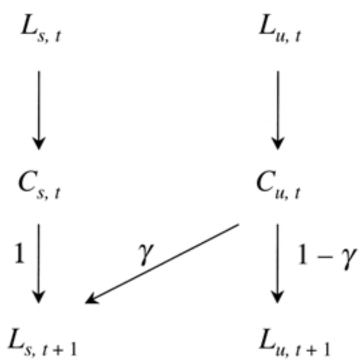

Figure 1. Dynamics of C_t^s, C_t^u, L_t^s and L_t^u.

We now consider the dynamics of the model. At time t, L_t^s is the number of skilled workers and L_t^u is the number of unskilled workers. Skilled and unskilled workers have n_t^s and n_t^u children respectively, who then choose whether or not to become educated at time $t+1$. Note that unskilled workers have more children than skilled workers. For the fraction of skilled to be constant, this fertility differential has to be offset by some children of unskilled becoming skilled (Figure 1 diagrams the dynamics of these variables). If all the children of skilled workers and a fraction γ_t of children of unskilled workers become skilled, then the evolution of R_t can be expressed as

$$R_{t+1} = \frac{L_{t+1}^s}{L_{t+1}^u} = \frac{L_t^s n_t^s + \gamma_t L_t^u n_t^u}{(1-\gamma_t) L_t^u n_t^u} = \frac{R_t^2 + \gamma_t}{1-\gamma_t}. \tag{4}$$

(From the above equation, it may appear that we assume all children of skilled workers become skilled workers, but in fact we assume only that at least a number $L_t^s n_t^s$ do so. An equilibrium exists only if $R_t \leq 1 - c^L$, and under this condition, it is a result that all children of skilled workers become skilled.) Given R_{t+1}, γ_t must satisfy

$$\begin{aligned} &\gamma_t \in [0,\theta] && \text{if } R_{t+1} = 1 - c^L, \\ &\gamma_t = \theta && \text{if } R_{t+1} \in (1-c^H, 1-c^L), \\ &\gamma_t \in [\theta, 1] && \text{if } R_{t+1} = 1 - c^H. \end{aligned} \tag{5}$$

2.2. *Steady States*

To solve for steady states, we look for fixed points of R and γ. Setting $R_{t+1} = R_t$ and $\gamma_{t+1} = \gamma_t$ in (4) implies that any steady state must satisfy the following quadratic equation:

$$\frac{1}{1-\gamma^*} R^{*2} - R^* + \frac{\gamma^*}{1-\gamma^*} = 0. \tag{6}$$

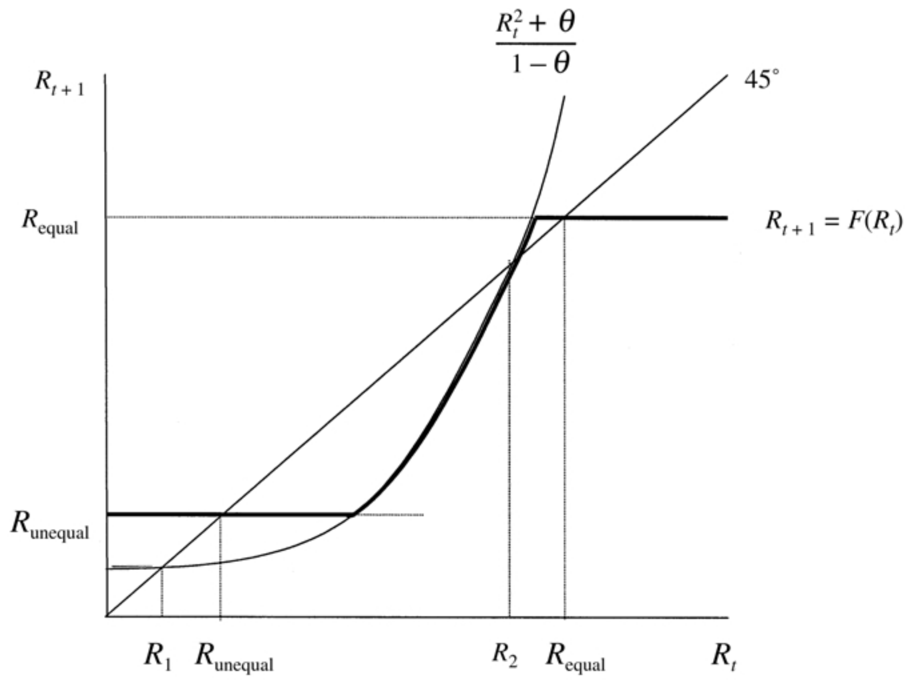

Figure 2. Multiple steady states.

Equations (5) and (6) have three possible types of solutions, corresponding to the three potential steady states shown in Figure 2 and the three sets of parameter values indicated in equation (5). (i) If the steady-state ratio of skilled to unskilled workers induces a wage differential of exactly $1/(1-c^L)$, then $R^* = R_{\text{equal}} = 1 - c^L$ and among children of unskilled workers, all those with a high cost of education will choose no education and some or all of those with a low cost of education will choose education (that is, $\gamma \leq \theta$). (ii) Similarly, if R^* induces a wage differential of $1/(1-c^H)$, then $R^* = R_{\text{unequal}} = 1 - c^H$ and $\gamma \geq \theta$. (iii) Finally, if the steady-state ratio of skilled to unskilled workers induces a wage differential between $1/(1-c^L)$ and $1/(1-c^H)$, then $\gamma = \theta$. All the children with a low cost of education will become educated and all the children with a high cost of education will not become educated. No other solutions are possible since the wage differential can never be expected to be outside the range $[1/(1-c^L),\ 1/(1-c^H)]$.

To see whether these potential steady states exist, note that at the equal steady state, R_{equal}, the unskilled have more children than the skilled, but enough children of the unskilled become skilled in each period to maintain the skilled–unskilled ratio at R_{equal} and the wage differential at $1/(1-c^L)$. Thus, the proportion θ of children of unskilled who have a low cost of education cannot be too small relative to the fertility differential between skilled and unskilled.

At the unequal steady state, the wage differential is $1/(1-c^H)$, and hence all children who have a low cost of education become skilled and some children with a high cost of education may do so as well. The unequal steady state is admissible if the proportion of children of unskilled workers with a low cost of education is low enough relative to the fertility differential between skilled and unskilled. In this case, even if all θ children become skilled, the ratio of skilled to unskilled workers in the next period does not increase and the wage differential does not fall below $1/(1-c^H)$.

Three conditions are needed for multiple steady states.

Theorem *Assumptions A1–A3 are necessary and sufficient conditions for multiple steady states.*

- *A1:* $0<\theta<3-2\sqrt{2}\equiv\theta_{\text{critical}}$.
- *A2: The time cost c^L is sufficiently low that $R_{\text{equal}}\in(R_2,1)$.*
- *A3: The time cost c^H is sufficiently high that $R_{\text{unequal}}\in(0,R_2)$.*

R_2 is the positive root of (6) with $\gamma_t=\theta$, where

$$R_2=[1-\theta+\sqrt{1-6\theta+\theta^2}]/2. \tag{7}$$

The proof is deferred to the appendix but the intuition behind these assumptions and the multiple steady states shown in Figure 2 is as follows. Assumption A1 assures R_1, the negative root of (6), and R_2 are real. Consider the curve $(R_t^2+\theta)/(1-\theta)$, which describes the evolution of R_t for $\gamma_t=\theta$. Under A1 this curve crosses the 45° line twice in the range $R_t\in(0,1)$. To understand Assumption A2 and A3, note that multiple steady states arise in this model only when children of the unskilled have varying costs of education (some children of the unskilled have a high cost of education, some have a low cost of education). Thus, multiple steady states will arise only if c^L and c^H (or alternatively the wage differentials at the equal and unequal steady states, $1/(1-c^L)$ and $1/(1-c^H)$, respectively), are sufficiently differentiated. Assumption A1 also fits this context. If θ is too high, then too many of the poor have a low cost of education, which makes them insufficiently differentiated for multiple steady states to arise.

What follows is a brief mathematical explanation of Figure 2, the complete analysis of which is deferred to the appendix. The evolution of R_t, denoted $R_{t+1}(R_t)$ and depicted in Figure 2, follows from (4) and (5) and is given by

$$\begin{aligned}
R_{t+1}&=R_{\text{unequal}} && \text{if } (R_t^2+\theta)/(1-\theta)\le R_{\text{unequal}};\\
R_{t+1}&=(R_t^2+\theta)/(1-\theta) && \text{if } (R_t^2+\theta)/(1-\theta)\in(R_{\text{unequal}},R_{\text{equal}});\\
R_{t+1}&=R_{\text{equal}} && \text{if } (R_t^2+\theta)/(1-\theta)\ge R_{\text{equal}}.
\end{aligned} \tag{8}$$

(1) If $(R_t^2+\theta)/(1-\theta)\le R_{\text{unequal}}$, then $R_{t+1}=R_{\text{unequal}}$. In this range $R_{t+1}>R_{\text{unequal}}$ is not an equilibrium. As follows from (5), for $R_{t+1}>R_{\text{unequal}}$, $\gamma_t=\theta$, implying that $R_{t+1}=(R_t^2+\theta)/(1-\theta)>R_{\text{unequal}}$, in contradiction. $R_t<R_{\text{unequal}}$ is not an equilibrium for any t, since all individuals would want to become skilled, thus it follows that if $(R_t^2+\theta)/(1-\theta)\le R_{\text{unequal}}$, then $R_{t+1}=R_{\text{unequal}}$.

(2) If $(R_t^2 + \theta)/(1 - \theta) \in (R_{\text{unequal}}, R_{\text{equal}})$, then $R_{t+1} = (R_t^2 + \theta)/(1 - \theta)$. In this range, $R_{t+1} = R_{\text{unequal}}$ is not an equilibrium, since for $R_{t+1} = R_{\text{unequal}}$, as follows from (5), $\gamma_t \geq \theta$, and hence $R_{t+1} \geq (R_t^2 + \theta)/(1 - \theta) > R_{\text{unequal}}$, in contradiction. In this range $R_{t+1} = R_{\text{equal}}$ is also not an equilibrium, since for $R_{t+1} = R_{\text{equal}}$ as follows from (5), $\gamma_t \leq \theta$ and hence $R_{t+1} \leq (R_t^2 + \theta)/(1 - \theta) < R_{\text{equal}}$, in contradiction. Hence, it follows that if $(R_t^2 + \theta)/(1 - \theta) \in (R_{\text{unequal}}, R_{\text{equal}})$, then $R_{t+1} \in (R_{\text{unequal}}, R_{\text{equal}})$ and, from (5), $\gamma_t = \theta$, therefore $R_{t+1} = (R_t^2 + \theta)/(1 - \theta)$.

(3) If $(R_t^2 + \theta)/(1 - \theta) \geq R_{\text{equal}}$, then by a similar argument as 1 above $R_{t+1} = R_{\text{equal}}$.

Under A1–A3 as depicted in Figure 2 there are two locally stable steady states at R_{unequal} and R_{equal}. R_2 is the unstable steady state—the threshold between the basin of attraction of the two stable steady states. If, however, $R_{\text{unequal}} < R_1$, then the high inequality steady state is R_1. If the low cost of schooling, c^L, is too high (A2 fails), then $R_{\text{equal}} = 1 - c^L < R_2$ and the system would have a unique steady state at R_{unequal}. If the high cost of schooling, c^H, is too low (A3 fails), then $R_{\text{unequal}} = 1 - c^H > R_2$ and the system would have a unique steady state at R_{equal}. If $\theta > 3 - 2\sqrt{2}$ (A1 fails), then the system would be characterized by a unique steady state R_{equal}.

Claim *All possible steady states can be second-order stochastically ranked by R^*, the ratio of skilled to unskilled workers.*

Proof. The Cobb–Douglas formulation fixes the share of income going to the educated and uneducated and is independent of the proportion of educated and uneducated. At the equal steady state, there are more educated workers than at the unequal steady state, so each educated worker necessarily gets a smaller share of income than what he would get at the unequal steady state. To get to the income distribution at the unequal steady state, income must necessarily be transferred from poorer people to richer people.[6] ■

2.3. Stability

Which steady state the economy converges to depends on initial conditions. If the initial ratio of unskilled to skilled, and hence the initial fertility differential between the unskilled and the skilled, is sufficiently great, then even if a proportion θ children of unskilled workers obtain education, the proportion of unskilled workers will rise in the next generation. In this case, the unskilled wage in the next generation will fall. The proportion of unskilled will keep growing until the wage differential between the groups becomes $1/(1 - c^H)$. On the other hand, if the initial ratio of unskilled to skilled, and hence the fertility differential between the unskilled and the skilled, is small enough, then the proportion of unskilled in the next generation will fall, and the economy will approach a steady state in which the wage differential is $1/(1 - c^L)$. R_2 is the critical population ratio that is on the borderline between these cases and is an unstable steady state.

2.4. Comparative Statics

In general, as the proportion of children with low cost of education increases or the cost of education falls, wage differentials decrease and the basin of attraction to the unequal steady state shrinks, possibly becoming eliminated. Please refer to the appendix (and Figure 3) for a more thorough discussion of the comparative dynamics.

To derive the comparative statics, we first consider how the basin of attraction of the stable steady states and the wage differentials at those steady states are affected by changes in the underlying parameters θ, c^L, and c^H; that is, we consider how R_{equal}, R_{unequal}, R_2, and R_1 themselves change as the underlying parameters change. We then discuss how changes in the same parameters affect the cutoff values for the admissibility of R_{equal}, R_{unequal}, R_2, and R_1.

When there are multiple steady states as in Figure 2, the system converges to the unequal steady state if R_0, the initial ratio of skilled to unskilled workers, is less than R_2, and to the equal steady state if R_0 is greater than R_2. The parameter that affects R_2 and thus the basin of attraction from which the system approaches either the unequal steady state or the equal steady state is θ. Increases in θ, the proportion of children of unskilled parents who have a low cost of education, reduce R_2, and expand the basin of attraction of the equal steady state. To see this, note that

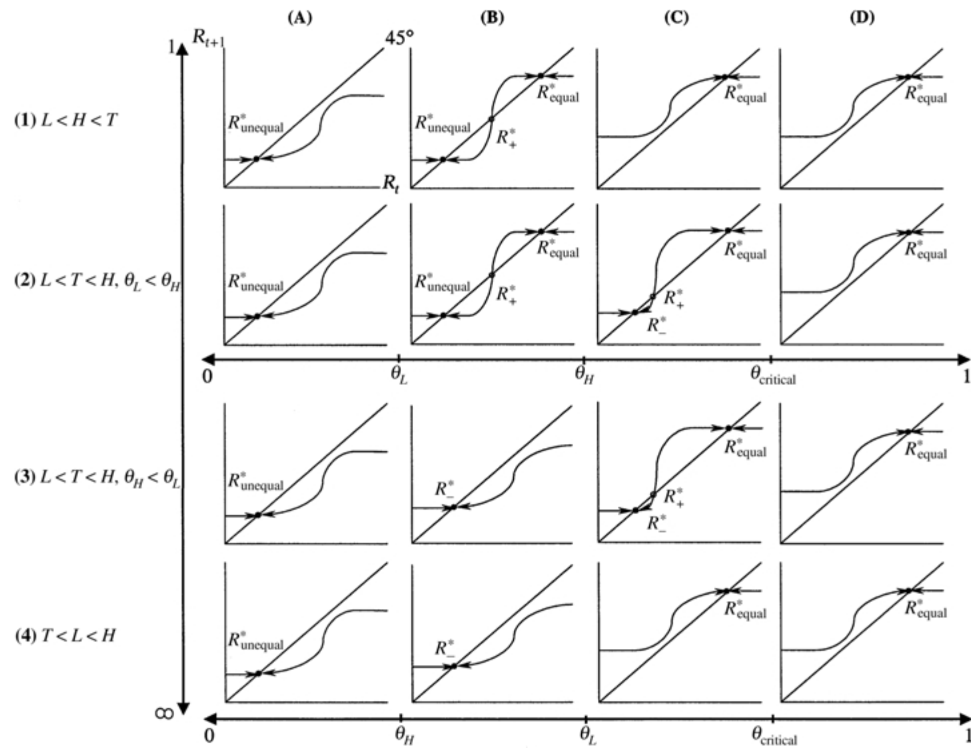

Figure 3. Dynamics of $R_{t+1}(R_t)$ as θ, L, and H vary.

$$\frac{\delta R_2}{\delta\theta} = \frac{-1}{2}\left(1 + \frac{(3-\theta)}{\sqrt{1-6\theta+\theta^2}}\right), \tag{9}$$

which is negative since θ is less than 1. Since R_2 exists, the term under the square root is positive.

The model thus suggests that countries with R_0 just under R_2 may face a brief window of opportunity in which small and temporary increases in θ can move them into the basin of attraction for the equal steady state. As time passes, and R falls, larger or longer-lasting increases in θ would be necessary to move to the more equal steady state.

If a country wishes to reduce the steady-state wage differential at its current steady state, the model suggests reducing c^L or c^H, the cost of education for different segments of the population, suffices to reduce the corresponding steady-state wage differential, $1/(1-c^L)$ and $1/(1-c^H)$, at the equal and unequal steady states. For some parameter values, however, R_1 replaces R_{unequal} as the unequal steady state, in which case changes of θ may also affect the steady-state wage differential. Increasing θ, the proportion of children of unskilled with low cost of education, results in a lower wage differential. To see this, refer back to the quadratic equation in (6). Note that increases in θ increase the coefficient of R^2 and the constant term. This means the upwards-pointing parabola narrows and moves strictly upwards and does not cross the original parabola. Thus if R_2, one root of the quadratic equation, decreases as θ increases, as shown in equation (9), then the other root, R_1, must increase, which implies θ and the wage differential at R_1 are inversely related.

In summary, our model suggests that if a country is at an unequal steady state, a temporary increase in the probability that children of unskilled parents become skilled, due, for example, to temporary expansion of educational opportunities, may induce a shift into the basin of attraction of the more equal steady state and permanently move the country to greater equality. This temporary increase must last long enough for the country

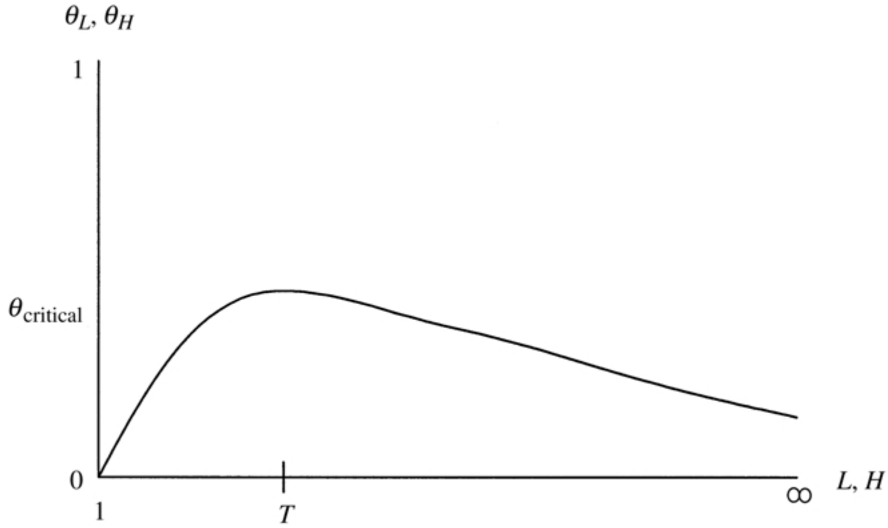

Figure 4. $\theta_L(L)$ and $\theta_H(H)$.

to move near the equal steady state so that when the push ends, the country is still in the basin of attraction of the equal steady state. For some parameter values, a temporary decrease in the cost of education for children with a high cost of education has a similar effect.

3. Empirical Evidence

We find some evidence that the fertility differential between educated and uneducated women is greater in countries with more inequality as the model implies.[7] Of course causality runs in both directions in the feedback model, and even if only one direction of the causality exists, there would still be a positive association between differential fertility and inequality.

The model predicts that there should be a positive relationship between differential fertility and Gini coefficients. Recall that the income distribution at the unequal steady state second-order stochastically dominates the income distribution at the equal steady state.

The remainder of the section is organized as follows. Section 3.1 discusses the sources for the data (Table 1 summarizes the available data).[8] Section 3.2 describes the methodology and empirical models. Section 3.3 reports the results.

3.1. Data

To measure differential fertility, we use data on total fertility rates (TFR) by women's educational attainment from four comparative studies (United Nations, 1987; Jones, 1982; United Nations, 1995; Mboup and Saha, 1998).[9] The four comparative studies calculate TFR by summing age-specific fertility rates. Thus, we can interpret the total fertility rate as the expected number of children a woman would have should she live until the end of her reproductive years and experience the age-specific fertility schedule. Table 2 reports the available data. Data on income inequality comes from Deininger and Squire (1996). Most of these Gini coefficients are calculated based on gross household income. GDP data

Table 1. Available data.

	Number of Observations	Mean	Standard Deviation
Fertility differentials vs. Gini coefficients			
Countries with fertility differentials	88		
Country-years with fertility differentials	96		
Fertility differentials	96	0.058	0.034
Gini coefficients	88	0.423	0.094
GDP	96	2,851	2,829

comes from Summers and Heston (1991), who report real GDP per capita in constant dollars (Chain index) expressed in international prices, base 1985.

3.2. *Methodology*

To measure differential fertility for each country, we regress ln(TFR) as reported in Table 2 on years of education. We weight observations in this regression by the percentage of women in the education category.[10] Weights are used because in some countries there are very few women with extreme levels of educational attainment, which increases the noise with which fertility is measured at these levels. We report the negative of the coefficient in this regression, so a positive number implies uneducated women have more children than educated women. We typically assume that all women in an education category have the average of the range of years of education in the category.

Table 2. TFR by women's years of education.

(A) World Fertility Survey, 1974–1982, Developing Countries (United Nations, 1987)

Country	Survey Year	Wtd OLS	Std Error	TFR for Education Category 0	1–3	4–6	7+	% of Women in Category 0	1–3	4–6	7–9	10+
Bangladesh	1975	−0.001	0.017	6.1	6.4	6.7	5.0	78	8	11	2	1
Benin	1982	0.056	0.019	7.4	8.5	5.8	4.3	88	3	6	2	1
Cameroon	1978	0.006	0.017	6.4	7.0	6.8	5.2	68	7	19	4	2
Colombia	1976	0.131	0.011	7.0	6.0	3.8	2.6	16	38	29	11	5
Costa Rica	1976	0.091	0.013	5.0	5.0	3.6	2.7	8	25	41	9	13
Cote D'Ivoire	1980	0.029	0.009	7.4	8.0	6.4	5.8	84	4	8	3	2
Dominican Rep.	1975	0.114	0.032	7.0	7.3	5.4	3.0	16	38	28	12	7
Ecuador	1979	0.138	0.031	7.8	7.2	5.3	2.7	14	25	35	11	14
Ghana	1979	0.025	0.006	6.8	6.7	6.7	5.5	60	3	7	9	20
Guyana	1977	0.048	0.007	6.6	7.0	5.6	4.8	4	3	23	60	6
Haiti	1977	0.089	0.009	6.0	4.8	4.1	2.8	70	15	9	3	2
Jamaica	1975	0.044	0.011	6.2	5.9	5.8	4.8	2	3	18	67	10
Jordan	1976	0.076	0.010	9.3	8.6	7.0	4.9	50	6	21	13	9
Kenya	1977	0.012	0.012	8.3	9.2	8.4	7.3	53	12	18	12	4
Korea, Rep.	1974	0.066	0.007	5.7	5.5	4.3	3.4	21	8	42	17	12
Lesotho	1977	0.029	0.021	6.2	5.6	6.0	4.8	8	13	55	23	2
Malaysia	1974	0.049	0.022	5.3	5.3	4.8	3.2	35	18	34	7	6
Mexico	1976	0.108	0.024	8.1	7.5	5.8	3.3	22	33	27	13	4
Morocco	1979	0.061	0.009	6.4	5.2	4.4	4.2	88	2	6	2	2
Pakistan	1975	0.068	0.028	6.5	5.4	6.1	3.1	87	2	5	2	2
Panama	1975	0.128	0.023	7.0	6.9	5.0	3.0	6	14	40	18	22
Paraguay	1979	0.132	0.008	8.2	6.6	4.6	2.9	7	31	41	9	12
Peru	1977	0.098	0.015	7.3	6.8	5.1	3.3	31	24	23	8	14
Philippines	1978	0.080	0.046	5.4	7.0	6.2	3.8	6	12	48	9	24
Senegal	1978	0.045	0.023	7.3	9.4	6.3	4.5	90	2	5	2	1
Sudan	1978	0.070	0.010	6.5	5.6	5.0	3.4	81	7	7	2	2
Syria	1978	0.094	0.005	8.8	6.7	5.6	4.1	67	4	18	7	5
Trinidad & Tobago	1977	0.047	0.018	4.6	3.4	4.1	3.2	4	2	12	67	15
Venezuela	1977	0.129	0.024	7.0	6.4	4.6	2.6	14	16	43	16	11
Yemen	1979	0.093		8.6	5.4	5.4	5.4	98	0	1	0	0
Averages	1977	0.072	0.017	6.9	6.5	5.5	4.0	42.5	12.6	22.7	14.0	7.7

Table 2. Continued.

(B) Demographic and Health Surveys, 1985–1989, Developing Countries (United Nations, 1995)

Country	Survey Year	Wtd OLS	Std Error	TFR for Education Category					% of Women in Category				
				0	1–3	4–6	7–9	10+	0	1–3	4–6	7–9	10+
Bolivia	1989	0.077	0.013	6.2	6.4	5.3	4.2	2.8	17.5	21.6	19.9	15.4	25.7
Botswana	1988	0.050	0.012	5.9	5.6	5.1	4.5	3.1	24.1	7.7	16.7	34.5	17.0
Brazil	1986	0.096	0.011	6.7	5.2	3.4	2.8	2.2	7.4	22.3	31.6	16.0	22.6
Burundi	1987	0.016	0.011	7.0	7.4	6.7	6.6	4.2	80.2	6.8	10.8	1.1	0.8
Colombia	1986	0.103	0.006	5.6	4.5	3.6	2.5	1.8	6.9	23.9	31.3	21.0	16.8
Dominican Rep.	1986	0.073	0.007	5.8	5.0	4.4	3.5	2.6	5.9	20.9	24.7	21.0	27.6
Ecuador	1987	0.092	0.007	6.4	6.3	4.7	3.5	2.6	7.8	14.8	32.7	16.1	28.6
Egypt	1988	0.057	0.006	5.7	5.3	4.2	3.4	3.4	52.0	13.5	17.3	11.4	5.9
El Salvador	1985	0.077	0.006	6.0	5.2	3.9	3.5	2.5	21.3	24.6	24.6	13.4	16.0
Ghana	1988	0.029	0.010	7.1	6.6	6.4	6.8	4.9	39.7	5.8	10.4	15.9	28.0
Guatemala	1987	0.093	0.008	6.9	5.6	4.2	2.8	2.7	41.7	24.1	19.6	6.2	8.4
Indonesia	1987	0.036	0.011	3.8	4.0	3.6	2.8	2.6	23.2	20.8	39.1	8.6	8.2
Kenya	1988	0.033	0.014	7.2	7.5	7.5	6.2	4.6	25.1	7.4	20.2	34.2	12.8
Liberia	1986	0.032	0.015	6.8	7.1	7.5	5.7	4.2	63.0	6.9	11.1	9.2	9.7
Mali	1987	0.025	0.005	7.0	6.7	6.6	5.7	4.7	85.4	3.4	6.3	3.8	1.1
Mexico	1987	0.106	0.014	6.4	6.3	4.0	2.7	2.4	11.6	16.6	31.4	26.4	14.0
Morocco	1987	0.099	0.012	5.5	3.9	2.9	2.4	2.2	82.7	3.6	7.1	3.1	3.3
Peru	1986	0.097	0.005	7.4	6.1	4.6	3.7	2.5	10.9	17.8	24.3	17.0	29.9
Senegal	1986	0.058	0.003	7.0	6.4	5.5	4.3	3.6	77.4	3.5	10.1	4.1	4.8
Sri Lanka	1987	0.008	0.004	2.8	3.0	2.9	2.7	2.7	11.1	12.6	23.6	29.6	23.0
Thailand	1987	0.075	0.007	3.5	2.8	2.5	2.1	1.5	9.7	5.0	70.9	2.9	11.5
Togo	1988	0.049	0.013	7.2	7.1	6.0	3.9	4.8	58.7	10.1	19.0	7.2	4.9
Trinidad & Tobago	1987	0.033	0.019	2.3	4.3	3.6	3.8	2.9	1.1	3.8	22.9	21.4	50.8
Tunisia	1988	0.065	0.004	5.1	4.7	3.7	2.8	2.6	56.7	7.6	23.5	5.2	7.0
Uganda	1988	0.020	0.008	7.7	7.4	7.0	7.2	5.3	39.8	18.0	25.8	13.2	5.2
Zimbabwe	1988	0.070	0.015	7.3	7.2	6.3	5.0	3.3	13.6	10.5	23.6	35.1	17.0
Averages	1987	0.060	0.010	6.0	5.7	4.9	4.0	3.2	33.6	12.8	23.0	15.1	15.4

(C) Demographic and Health Surveys, 1990–1994, Developing Countries (Mboup, 1998)

Country	Survey Year	Wtd OLS	Std Error	TFR for Education Category			Years in Primary
				No School	Primary	Secondary+	
Bangladesh	1993	0.048	0.024	3.64	3.24	2.49	5
Bolivia	1993	0.064	0.060	5.84	5.62	3.05	6
Burkina Faso	1993	0.080	0.049	6.06	4.97	2.86	6
Cameroon	1991	0.020	0.029	5.78	6.02	4.44	7
Central Afr. Rep.	1994	0.021	0.029	4.73	4.94	3.74	6
Colombia	1990	0.097	0.030	4.66	3.36	2.24	5
Dominican Rep.	1991	0.049	0.019	4.80	3.71	2.75	8
Egypt	1992	0.058	0.013	4.82	3.66	2.94	6
Ghana	1993	0.055	0.040	5.36	4.56	2.69	8
Indonesia	1994	0.022	0.024	3.12	3.14	2.47	6
Jordan	1990	0.032	0.013	6.31	5.62	4.74	6
Kenya	1993	0.032	0.026	5.40	5.08	3.75	7
Madagascar	1992	0.030	0.041	5.88	6.21	4.19	6
Malawi	1992	0.041	0.034	6.11	5.81	4.08	6
Morocco	1992	0.119	0.013	4.16	2.15	1.85	5
Namibia	1992	0.041	0.030	5.64	5.26	3.80	6
Niger	1992	0.048	0.048	6.42	6.37	3.92	6
Nigeria	1990	0.036	0.037	5.67	5.66	3.91	6
Pakistan	1990	0.034	0.013	4.69	4.23	3.61	5

Table 2. Continued.

(C) Demographic and Health Surveys, 1990–1994, Developing Countries (Mboup, 1998)

				TFR for Education Category			
Country	Survey Year	Wtd OLS	Std Error	No School	Primary	Secondary +	Years in primary
Paraguay	1990	0.075	0.045	6.06	5.03	3.00	6
Peru	1991	0.123	0.060	6.49	4.80	2.48	5
Philippines	1993	0.020	0.051	4.31	5.13	3.25	6
Rwanda	1992	0.049	0.028	5.98	5.25	3.78	6
Senegal	1992	0.066	0.034	5.85	4.79	3.19	6
Turkey	1993	0.106	0.015	4.00	2.30	1.65	6
Zambia	1992	0.029	0.027	6.39	6.18	4.55	7
Zimbabwe	1994	0.032	0.024	4.68	4.36	3.28	7
Averages	1992	0.053	0.032	5.29	4.72	3.29	6

(D) World Fertility Survey, 1975–1979, Europe/USA (Jones, 1982)

				TFR for Education Category				
Country	Survey Year	Wtd OLS	Std Error	Elementary Incomplete	Elementary Complete	Low Secondary	High Secondary	Post-Secondary
Bulgaria	1976	0.046	0.010	2.41	1.74	1.55	1.50	1.37
Czechoslovakia	1977	0.041	0.004	2.35		2.08	1.80	1.62
Denmark	1975	0.019	0.007	2.20		1.87	1.86	1.85
Finland	1977	0.022	0.002	2.01		1.80	1.74	1.64
France	1977	0.034	0.006	2.51	2.03	1.86	1.79	1.66
Italy	1979	0.042	0.006	2.45	1.96	1.74	1.65	1.48
Norway	1977	0.028	0.001	2.40		2.11	1.95	1.86
Poland	1977	0.051	0.001	2.70	2.32	1.95	1.71	1.55
Romania	1978	0.053	0.007	2.25		1.68	1.52	1.39
Spain	1977	0.007	0.007	2.63	2.28	2.42	2.27	2.41
U.K.	1976	0.026	0.004	2.15		1.90	1.73	1.72
U.S.A.	1976	0.045	0.003	2.76		2.34	2.07	1.82
Yugoslavia	1976	0.045	0.010	2.43	1.81	1.57	1.57	1.40
Averages	1977	0.035	0.005	2.40	2.02	1.91	1.78	1.67

Notes: Data is taken from World Fertility Survey (United Nations, 1987; Jones, 1982) and Demographic and Health Surveys (United Nations, 1995; Mboup, 1998). For the last two groups of surveys, no data on percentage distribution of women is available. Weighted OLS coefficients are calculated by regressing TFR on proxy years representing the education categories and using percentage of Women in category as weights where available. If not available, percentage of women is assumed as uniform. Standard errors are from the OLS regression. Proxy years are: 0, 2, 5, and 8 for non-European/USA surveys from 1974–1982; 0, 2, 5, 8, and 11 for 1985–1989 surveys; 0, years in primary, primary + 2 for 1990–1994 surveys; 3, 6, 9, 12, and 14 for European surveys with 5 education categories; 5, 9, 12, and 14 for European/USA surveys with four education categories.

We test the model using data on total fertility rates by women's educational attainment and Gini coefficients.[11] TFR by women's educational attainment and Gini coefficients of inequality could be obtained for 88 observations in 62 countries from 1974 to 1994.

We use a model that specifies the TFR by women's educational attainment as independent across observations from different countries but not necessarily independent for observations within countries. Because the data covers three decades, the data set includes some countries up to three times. Since income inequality is likely to be correlated for observations of the same country due to unobservable characteristics of the environment, failing to account for these correlations would most likely lead to underestimation of the standard errors of the coefficients. More specifically, the model is:

$$F_{ct} = a + bG_{ct} + cX_{ct} + u_c + \varepsilon_{ct}, \qquad \text{Model (1)}$$

where c indexes countries and t indexes time. F_{ct} is the fertility differential, which is approximated as the coefficient from the weighted least squares regression described earlier; a is a constant; b reflects the relationship between the fertility differential and the Gini coefficient; G_{ct} is the Gini coefficient; X_{ct} is a vector of country and time variables such as ln(GDP), continent dummy variables, and the year of survey; u_c is a country-specific error term; and ε_{ct} is the residual.

We also take into account the fact that ε_{ct} is heteroskedastic. The noise in the approximation of F_{ct}, the fertility differential, varies across observations. The earlier regressions that provide the measurements of differential fertility, however, also provide an estimate of their noise—the standard errors of the same regression. In this case, where the standard errors are given by h_{it} and the variance of the error term is given by $\text{var}(\varepsilon_{ct}) = h_{it}^2\sigma^2$, we define $F_{ct}^* = F_{ct}/h_{it}$, $G_{ct}^* = G_{ct}/h_{it}$, $X_{ct}^* = X_{ct}/h_{it}$, $1^* = 1/h_{it}$, and $\varepsilon_{ct}^* = \varepsilon_{ct}/h_{it}$ and we estimate the model:

$$F_{ct}^* = a1^* + bG_{ct}^* + cX_{ct}^* + \varepsilon_{ct}^*. \qquad \text{Model (2)}$$

In effect, this regression puts more weight on observations with low variance.

We report the results from using (i) only clustering by country (model (1)), (ii) only weights (model (2)), and (iii) both clustering by country and weights (combination of models (1) and (2)). Since under the Kuznets hypothesis, Gini coefficients may be higher among middle-income countries than among either low-income or high-income countries, we control for linear and nonlinear income effects by adding ln(GDP) and $\ln(\text{GDP})^2$ as additional covariates.

3.3. Results

For most specifications, differential fertility seems positively correlated with Gini coefficients. Table 3 and Chart 1 refer to the sample of countries for which observations on TFR by women's educational attainment and Gini coefficients of inequality could be obtained.

Fertility differentials are significantly and positively correlated with Gini coefficients under the cluster and weighted specifications with or without controls for log per capita income and its square. Using both clustering by country and weights to correct for heteroskedasticity, the coefficient of the fertility differential regressed on the Gini coefficient is 0.089 and significant at the 1 percent level (see column (ix)), once one controls for the level and square of log per capita income. This indicates that going from a relatively equal country like Indonesia with a Gini coefficient of 0.320 in 1987 to a relatively unequal country like Brazil with a Gini coefficient of 0.545 in 1986 increases the fertility differential by 0.020. A fertility differential increase of 0.020 means, for example, the ratio of the expected number of children between a woman with no schooling and a

Table 3. Fertility differentials regressed on Gini coefficient, 1974–1995.

	Dependent variable: Weighted OLS coefficient of TFR on years of education								
	Cluster[1]			Weights[2]			Both[3]		
Independent variable	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Gini coefficient	0.101**	0.114***	0.093***	0.030	0.035	0.089***	0.030	0.035	0.089***
	(0.044)	(0.039)	(0.033)	(0.031)	(0.030)	(0.027)	(0.056)	(0.048)	(0.031)
Year	− 0.001	0.0001	0.0001	0.002	0.0004	0.0003	0.002	0.0005	0.0003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ln(GDP)		0.012**	0.311***		− 0.009**	0.289***		− 0.009	0.289***
		(0.006)	(0.063)		(0.004)	(0.050)		(0.005)	(0.073)
ln (GDP)^2			− 0.019***			− 0.018***			− 0.018***
			(0.004)			(0.003)			(0.005)
Constant	1.227	− 0.246	− 1.527	− 3.150	− 0.861	− 1.195	− 3.150	− 0.861	− 1.195
	(1.007)	(1.218)	(1.247)	(1.261)	(1.595)	(1.339)	(1.898)	(1.728)	(1.537)
N	88	88	88	88	88	88	88	88	88

Notes: GDP is real GDP per capita in constant dollars (Chain index) expressed in international prices, base 1985.
[1]These regressions correct standard errors for random effects (grouped errors by country).
[2]These regressions use weighted least squares (to weight for the precision of the first stage estimates).
[3]These regressions correct standard errors for random effects (grouped errors by country) and use weighted least squares (to weight for the precision of the first stage estimates).
Standard errors are in parentheses.
*Significant at 10 percent level.
**Significant at 5 percent level.
***Significant at 1 percent level.

woman with ten years of schooling is 1.22 times children greater in a country like Brazil than it is in a country like Indonesia.[12] Another interpretation is that increasing by two standard deviations of the Gini coefficient (two standard deviations is 0.188) increases the fertility differential by 0.017. Or, going from the 25th percentile of the Gini coefficients, 0.349, to the 75th percentile, 0.490, increases the fertility differential by 0.013. (Of course, these associations are not necessarily causal.)

There is some evidence that fertility differentials may be greatest among middle-income countries. The coefficient on ln(GDP) in column (ix) of Table 3 is positive and significant at the 1 percent level, while the coefficient on ln (GDP)^2 is negative and significant at the 1 percent level; for the range of income levels in this data set, fertility differentials initially rise among low to middle-income countries and then fall among middle to high-income countries, all else equal.

The results are reasonably robust to adding continent dummies. When we include a continent dummy for Latin America, the relationship between fertility differentials and the Gini coefficient is positive, but statistically insignificant; but when we include continent dummies for Latin America, sub-Saharan Africa, and Asia, the relationship is statistically significant at the 10 percent level.[13]

Ideally we would have wanted to use an instrument to shed light on the direction of causality. But in the absence of a good instrument, it is interesting to consider the example of Taiwan. During the communist takeover of China, Taiwan had exogenous increases in the proportion of skilled workers due to the exodus of educated people fleeing China. Since then, Taiwan has had a fairly equal wage distribution. Moreover, in Taiwan, there is

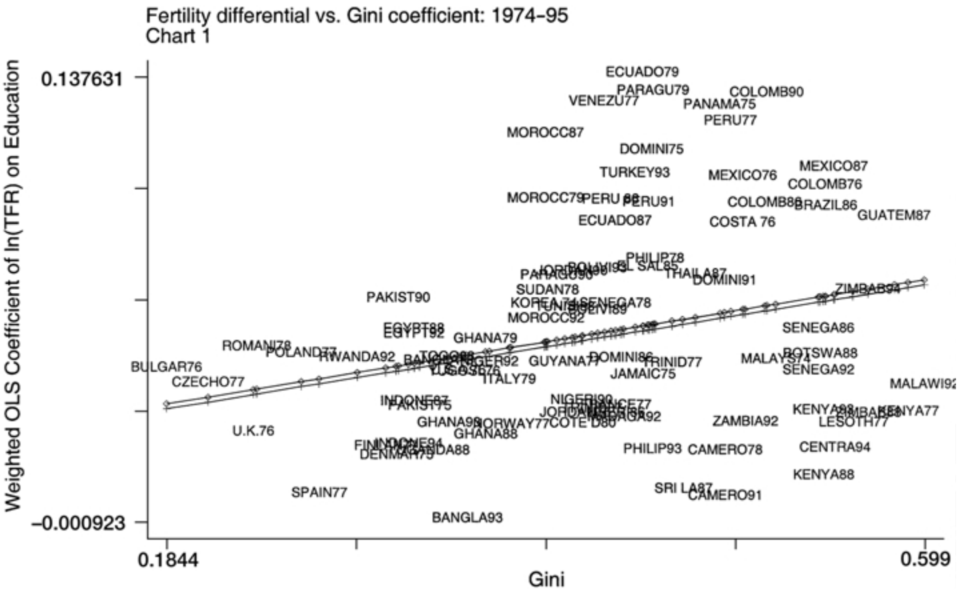

Chart 1. Fertility differential vs. Gini coefficient: 1974–1995.

no tendency for less educated women to have more children. In fact, women with some post-secondary education actually have twice as many children as women with primary school education or less[14] (Nei cheng pu, 1996).

4. A More General Relationship Between Fertility and Wages

In Sections 2 and 3, we assumed that fertility is inversely proportional to wages. We argue in this section that this may be an acceptable approximation over the moderate wage levels characteristic of middle-income countries, but it is not likely to fit as well at very low or very high wages as discussed below. We consider the more general relationship between fertility and wages and argue that the positive relationship between fertility differentials and income inequality—and hence the phenomenon of multiple equilibria emphasized in the model we presented in Section 2—is most likely for middle-income countries. Middle-income countries can be thought of as those with intermediate values of the productivity parameter A in our model. Thus, the generalization of the model is that, at a coarse grain, A determines national incomes, and at a finer grain, feedback between fertility and wages determines income levels, but mainly for middle-income countries. This section tests the prediction of this expanded model: we show that the relationship between fertility differentials and income inequality is most pronounced in middle-income countries.

At very low wages, wage increases may increase the number of surviving children by reducing infant mortality and infertility due to disease and malnutrition. In fact, among the

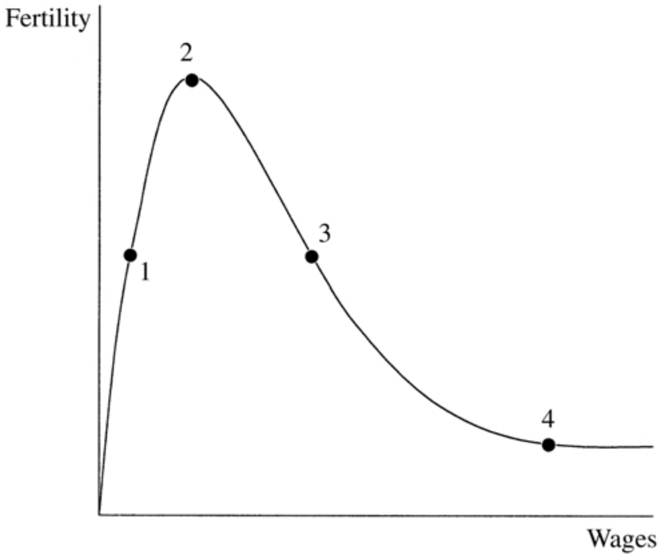

Figure 5. Fertility vs. wages.

poorest African countries, fertility initially increases with education. Child survival is also more important to model in the poorest countries. In middle-income countries, the under-5 mortality rate in 1997 was only 42 per 1,000 (World Bank, 1999), so it may be acceptable to neglect infant mortality, as we have done so far, and approximate differences in the number of surviving children of educated and uneducated workers by differences in fertility rates between these groups. The under-5 mortality rate was higher in low-income countries at 97 per 1,000 in 1997 (excluding China and India, the rate was 130 per 1,000; World Bank, 1999). Thus, in the very poorest countries, a high proportion of children of the unskilled may not survive, so increases in wages may have a strong positive effect on survival rates. Barro and Sala-i-Martin (1995) find that among the very poorest countries, fertility rises with income. At very high wages, further wage increases are likely to reduce fertility only modestly. This suggests that the relationship between fertility differentials and income inequality would be greatest among middle-income countries, negative among low-income countries, and weak among high-income countries.[15]

The relationship between the number of surviving children and wages is shown in Figure 5. At very low wages, the effect of inequality on fertility differentials depends on the general level of wages, and hence on the productivity parameter A. If wages are in region 1, then skilled workers have more children than unskilled workers. If wage differentials increase, then the fertility differential, as defined in Section 2, falls. If a low-income country is in region 2, then the relationship between inequality and fertility differentials is ambiguous.

The model applies best for middle-income countries around region 3. For these countries, fertility differentials, as measured in Section 3, should be the greatest relative to

Table 4. Fertility differentials regressed on Gini coefficient, 1974–1995.

	Dependent variable: Weighted OLS coefficient of TFR on years of education								
	Cluster[1]			Weights[2]			Both[3]		
Independent variable	Low (i)	Middle (ii)	High (iii)	Low (iv)	Middle (v)	High (vi)	Low (vii)	Middle (viii)	High (ix)
Gini coefficient	– 0.034	0.181***	0.267	0.045	0.147***	0.088	0.045	0.147***	0.088*
	(0.028)	(0.038)	(0.159)	(0.031)	(0.040)	(0.065)	(0.051)	(0.042)	(0.046)
Year	0.0004	– 0.001	– 0.002	0.001	– 0.001	– 0.002	0.0004	– 0.001	– 0.002
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)	(0.004)
Constant	– 0.827	1.696	4.413	– 0.771	1.690	4.199	– 0.771	1.690	4.199
	(1.140)	(1.442)	(3.694)	(1.447)	(1.651)	(7.341)	(1.626)	(1.768)	(7.467)
N	31	46	11	31	46	11	31	46	11

Notes: Low-income countries: with GDP measure < 1640 US$. Middle-income countries: with GDP measure between 1640 and 7169 US$. High-income countries: with GDP measure > 7169 US$.
[1]These regressions correct standard errors for random effects (grouped errors by country).
[2]These regressions use weighted least squares (to weight for the precision of the first stage estimates).
[3]These regressions correct standard errors for random effects (grouped errors by country) and use weighted least squares (to weight for the precision of the first stage estimates).
GDP is real GDP per capita in constant dollars (Chain index) expressed in international prices, base 1985. Standard errors are in parentheses.
*Significant at 10% level.
**Significant at 5% level.
***Significant at 1% level.

low and high-income countries. Consequently, the relationship between differential fertility and inequality should also be the strongest.

Among high-income countries, the fertility of both skilled and unskilled labor is small as shown in region 4. Since fertility asymptotes to a non-zero level, fertility differentials are small. Hence, measurement noise may obscure the relationship between differential fertility and inequality.

We find that the positive relationship between inequality and differential fertility is primarily found in the sample consisting of middle-income countries, while among high-income countries the relationship is weaker, and among the low-income countries the relationship is weakest (See Table 4.) Following the World Bank, we take the cut-off between middle-income and low-income developing countries as $1,640 GDP per capita while the cut-off between middle-income and high-income countries as $7,170 GDP per capita.[16]

Among middle-income countries, the association between differential fertility and the Gini coefficient has a larger magnitude than the coefficient for the sample of all countries and is statistically significant at the 5 percent level. The number 0.147 in column (ii) can be interpreted as increasing the fertility differential by 0.033 when inequality increases from the 1987 Indonesian level (with a Gini coefficient of 0.320) to the 1986 Brazilian level (with a Gini coefficient of 0.545). A fertility differential increase of 0.033 means, for example, the ratio of the expected number of children of a woman with no schooling to a woman with 10 years of education is 1.39 times greater in a country like Brazil than it is in a country like Indonesia.

For low-income countries the association between differential fertility and Gini coefficient is close to zero. Among high-income countries, the association between differential fertility and Gini coefficients is positive, though weaker than among middle-income countries in most specifications. Note, however, that the strength of this association in rich countries is mainly driven by the oil-rich outliers, Venezuela and Trinidad and Tobago, in this sub-sample of mainly European and North American countries. Indeed, the cut-off values for middle-income and high-income countries are arbitrary. If we use cut-off values of \$1,000 and \$9,500 GDP per capita for middle and high-income countries, then the relationship between differential fertility and Gini coefficients in high-income countries is not significant, and the relationship among low-income countries is negative and statistically significant for one specification.

Note that given A, an economy at the equal steady state has more skilled workers than an economy at the unequal steady state. Hence the more equal economy should also be richer. The model then suggests that there will be a positive correlation between income and equality.

5. Conclusion

The large fertility differentials between educated and uneducated workers found in some developing countries, together with intergenerational persistence in education levels, make it difficult to reduce inequality in these countries. Differential fertility tends to increase the proportion of unskilled workers, reducing their wages. Since wage reductions lower the opportunity cost of having children, there may be a positive feedback. A model incorporating this effect generates multiple steady-states with varying levels of inequality.

Consistent with the model, we find positive relationships between Gini coefficients and fertility differentials between educated women and uneducated women. The relationship between differential fertility and inequality is strongest for middle-income countries, weaker for the richest group of countries, and negative for the poorest countries.

Our model suggests that if a country is in an unequal steady state, a temporary increase in the probability that children of unskilled parents become skilled, due, for example, to an expansion of access to educational opportunities, may permanently move the country to a more equal steady state.

The analysis can be generalized and extended in several ways.[17] In the model, we assume a two-point form for the distribution of the cost of education. This discrete form of the distribution of cost implies that the supply of skilled labor is infinitely elastic at wage premiums of $1/(1-c^L)$ and $1/(1-c^H)$ and completely inelastic for wage premiums in between. If the wage premium ever falls below $1/(1-c^L)$, no children would invest in education; if the wage premium rises above $1/(1-c^L)$, all children of skilled workers and a fraction θ of children of unskilled workers would invest in education. In between $1/(1-c^L)$ and $1/(1-c^H)$, the ratio of children of unskilled workers who become skilled is θ. Consequently, between $1/(1-c^L)$ and $1/(1-c^H)$, the demographic force for instability explored in this paper dominates any effect of wage premia on incentives for education. Once the wage premium reaches $1/(1-c^L)$ or $1/(1-c^H)$, the traditional, stabilizing force by which increasing wage differentials increase incentives for education dominates. With a more general, continuous distribution of cost of education, either force could locally dominate in different areas. We conjecture that, with a more general distribution of cost of education, the R_{t+1} (R_t) curve could cross the 45° line an arbitrary number of times, generating an arbitrary number of stable steady states.

With a continuous distribution of cost, endogenous fertility has an additional effect that is suppressed under the discrete distribution of cost used in our model: the responsiveness of fertility to wage differentials makes the reaction function have a positive slope around a steady state. With discrete costs, whether fertility is exogenous or endogenous, the response curve is flat around the stable steady states (since the supply of skilled labor is infinitely elastic as an artifact of rational expectations).[18] With continuous costs, this remains true if fertility is exogenous. However, if fertility is endogenous, the response of fertility differentials to wages creates a positive feedback that should make R_{t+1} increase more steeply in R_t than if fertility were exogenous. Consider a small increase in the proportion of unskilled workers. This reduces the wages of unskilled workers. Under endogenous fertility, this increases their fertility, and hence the proportion of unskilled workers in the next generation will be larger than under exogenous fertility. Thus, the reaction function is steeper. One implication is that some policy interventions could have a larger multiplier effect with endogenous fertility. For example, suppose a public program educates an additional 1,000,000 children in each generation. Unskilled wages will rise, raising the opportunity cost of childbearing among the unskilled and reducing their fertility, which will further increase wages among the unskilled, creating a multiplier effect on inequality. Figure 6 shows, in a local neighborhood of a stable steady state, that a permanent increase in the function $R_{t+1}(R_t)$ leads to a larger steady-state proportion of skilled workers when there is endogenous fertility than when there is not.

If inequality is derived primarily by differences in ownership of capital and land, we

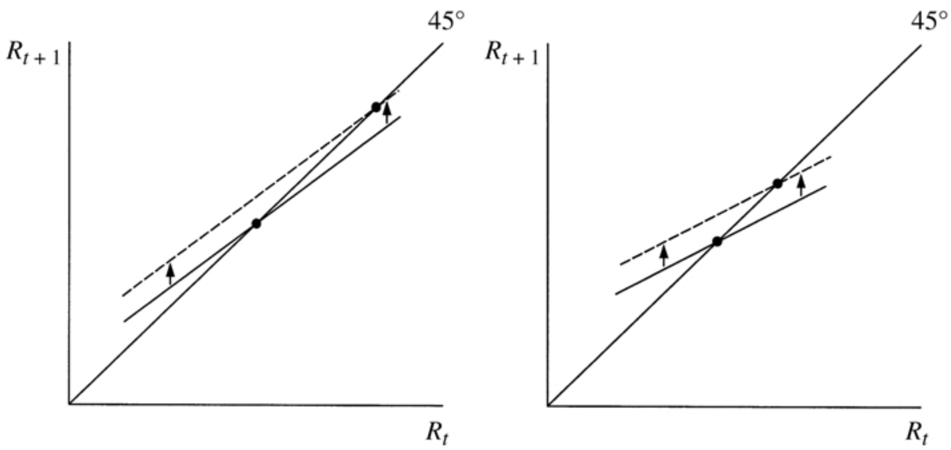

Figure 6. Multiplier effect in a local neighborhood of a steady state with and without endogenous fertility.

conjecture a similar story could be told. Suppose there are few land owners and many workers. If the number of workers is increased, this drives down wages, so workers have more children. Land owners will have more income and hence more children due to income effects. With additional population in the next generation, the factor distribution of income will tilt towards land. The impact on the Gini coefficient will depend on parameter values and inheritance rules, but it is likely that the Gini coefficient will rise.

If we consider men and women separately, the income effect is likely to be more applicable to men while the substitution effect is likely to be more applicable to women. It is an empirical question which one dominates at the household level. If income effect dominates, then our story may not apply. However, the empirical evidence we review is consistent with the hypothesis that substitution effects dominate overall.

We focus on a particular fertility function but conjecture that the results in this paper hold more generally as long as (i) skilled and unskilled workers are complements in production, (ii) children of the unskilled are more likely to be unskilled, and (iii) higher wages reduce fertility.

Appendix

Section A of this Appendix fully derives the admissibility of the various steady states and proves why Assumptions A1–A3 are necessary and sufficient conditions for multiple steady states. Section B explains why changes in certain parameter values may imply qualitative changes in the dynamical system in that steady state and threshold values change or disappear.

A. Admissibility

Define a steady state as the pair (R^*,γ^*), such that the ratio of skilled to unskilled workers at time t is R^*. If fertility and education decisions are taken optimally, the proportion of children of unskilled workers who become skilled will be γ^*, and the ratio of skilled to unskilled workers in the next generation will remain at R^*.

Recall that the evolution of R_t is given by,

$$R_{t+1} = \frac{R_t^2 + \gamma_t}{1 - \gamma_t}, \tag{10}$$

and that γ_t is determined in equilibrium given R_{t+1}, by

$$\begin{aligned} &\gamma_t \in [0, \theta] && \text{if } R_{t+1} = R_{\text{equal}}; \\ &\gamma_t = \theta && \text{if } R_{t+1} \in (R_{\text{unequal}}, R_{\text{equal}}); \\ &\gamma_t \in [\theta, 1] && \text{if } R_{t+1} = R_{\text{unequal}}. \end{aligned} \tag{11}$$

Setting $R_{t+1} = R_t$ in (10) implies that any steady state must satisfy the following quadratic equation:

$$\frac{1}{1-\gamma^*}R^{*2} - R^* + \frac{\gamma^*}{1-\gamma^*} = 0. \tag{12}$$

For ease of exposition, define $L = 1/(1-c^L)$ and $H = 1/(1-c^H)$.

Proposition 1 *The following is an admissible steady state*

$$[R_{\text{equal}},\ \gamma_{\text{equal}}] = \left[\frac{1}{L},\ \frac{L-1}{L^2+L}\right], \tag{13}$$

if and only if

$$\theta \geq \frac{L-1}{L^2+L} \equiv \theta_L. \tag{14}$$

Proof. Substituting R_{equal} into (12) yields the expression for γ_{equal} (13). Recall $1 - c^L = 1/L$. The condition $R = 1 - c^L$ is equivalent to saying that the children of unskilled workers with a low cost of education are indifferent to obtaining education. Since these children are represented by the fraction θ, this last statement, in turn, is equivalent to saying $\gamma_{\text{equal}} \leq \theta$, which is the inequality in (14). ■

Proposition 2 *The following is an admissible steady state*

$$[R_{\text{unequal}}, \gamma_{\text{unequal}}] = \left[\frac{1}{H}, \frac{H-1}{H^2+H}\right], \tag{15}$$

if and only if

$$\theta \leq \frac{H-1}{H^2+H} \equiv \theta_H. \tag{16}$$

Proof. The proof is similar to that of Proposition 1 except H replaces L. ■

The last case to consider is when the wage differential is between the low and high cost of education, in which case children obtain education if and only if they have a low cost of doing so.

Proposition 3 *R_i is a steady state only if the corresponding wage differential is between L and H, for i = 1, 2, where*

$$R_1 = [1 - \theta - \sqrt{1 - 6\theta + \theta^2}]/2,$$
$$R_2 = [1 - \theta + \sqrt{1 - 6\theta + \theta^2}]/2. \tag{17}$$

The wage differential at R_1 is at least as high as the wage differential at R_2.

Proof. Wage differentials are between L and H if and only if $\gamma^* = \theta$ (from equation 11). Equation (17) follows immediately from solving for the roots of (12) and substituting $\gamma^* = \theta$. The wage differential at R_1 is greater than the wage differential at R_2 since wage differentials are the inverse of R and $R_2 \geq R_1$. ■

Proposition 3 leads immediately to the following lemma, where we derive the value of θ for which $R_1 = R_2$.

Lemma 1 $R_1 = R_2$ if and only if $\theta = \theta_{\text{critical}}$, where $\theta_{\text{critical}} = 3 - 2\sqrt{2}$. If $\theta > \theta_{\text{critical}}$, R_2 and R_1 do not exist. If $\theta < \theta_{\text{critical}}$, R_1 and R_2 exist and R_1 does not equal R_2.

Proof. $R_1 = R_2$ if and only if the term under the square root in (17) is zero, which simplifies to a quadratic equation in θ and can be solved. The positive root is greater than 1, thus too large, while the negative root is admissible. The negative root is θ_{critical}, the value of θ for which $R_1 = R_2$. If $\theta > \theta_{\text{critical}}$, the term under the square root in (17) is negative. Hence neither R_1 nor R_2 exist. If $\theta < \theta_{\text{critical}}$, the term under the square root in (17) is positive, so R_1 and R_2 must be different. ■

Now we discuss how changes in the underlying parameters affect the admissibility of R_{equal}, R_{unequal}, R_1, and R_2. Figure 3 shows how the dynamics of $R_{t+1}(R_t)$ depend on the values of θ, L, and H (it will be helpful to refer to Figure 3 frequently). Since it seems plausible that improving the education system or subsidizing education could increase θ or decrease L and H, it is useful to discuss how the dynamics depend on θ, L, and H. On the two X-axes, θ varies from 0 to 1. On the Y-axis, L and H vary from 1 to infinity. The axes are discussed below. First note the following lemma.

Lemma 2

i. θ_L and θ_H are strictly between 0 and 1.

ii. θ_L and θ_H attain their maximum, θ_{critical}, when L and $H = T$, where $T = 1 + \sqrt{2}$. θ_L and θ_H have no other local maxima.

iii. $\theta_L < \theta_H$ implies $L < T$. $\theta_H < \theta_L$ implies $H > T$.

Proof. The proof follows from straightforward algebraic manipulation and is omitted.[19] ■

Figure 4 shows $\theta_L(L)$ and is helpful to understand the axes in Figure 3. On Row 1 in Figure 3, $L < H < T$ so $\theta_L < \theta_H$. On Row 2, $L < T < H$ but H is sufficiently close to T such that $\theta_L < \theta_H$. On Row 3, $L < T < H$ but this time L is sufficiently close to T such that $\theta_H < \theta_L$. On Row 4, $T < L < H$ so $\theta_H < \theta_L$.

The intuition for Figure 3 is as follows. The intuition is clearest for the lowest and highest values of θ (Columns A and D). If the proportion of children of unskilled with a low cost of education is very low, the economy converges to R_{unequal} (proved in Proposition 4). Recall the argument that for R_{equal} to be admissible, the proportion of children of unskilled who have a low cost of education cannot be too small relative to the fertility differential between skilled and unskilled. If the proportion of children of unskilled who have a low cost of education is too small, then too few children of unskilled will become skilled and the proportion of skilled adults in the population will shrink. Thus, there cannot be a steady state where the equilibrium wage differential is L.

For very high proportions of children of unskilled with a low cost of education, the economy converges to R_{equal}. R_{unequal} is inadmissible when the proportion of children of unskilled who have a low cost of education is too high relative to the fertility differential between skilled and unskilled. If θ is greater than θ_{critical}, then sufficiently many children of unskilled will become skilled so that the proportion of skilled adults in the population will rise.

Figure 3 broadly confirms the intuition that lower costs of education and a higher proportion of children of unskilled with a low cost of education may help reduce inequality. As the cost of education falls, the economy moves up along the Y-axis towards Row 1. As the proportion of children of unskilled parents with low cost of education increases, the economy moves right along the X-axis. As both the cost of education falls and the proportion of children of unskilled with low cost of education increases, the economy goes from a single unequal steady state to multiple steady states to a single equal steady state. The lowest values of L and H (Row 1) have the largest region of θ over which only the equal steady state is admissible (Columns C and D). As H increases beyond T (Row 2), this region shrinks (Column D). For even larger values of L and H (Row 3), the region over which only an unequal steady state is admissible increases (Columns A and B). The largest values of L and H (Row 4) have one steady state for all values of θ but the wage differentials at all the steady states are high.

To prove the relationships shown in Figure 3, we need the following lemma:

Lemma 3.

i. If $\theta < \theta_H$, then $R_1 < R_{\text{unequal}}$.[20]

ii. If $\theta < \theta_H$, then $R_2 > R_{\text{unequal}}$.

iii. If $\theta_H < \theta < \theta_{\text{critical}}$ and $H > T$, then $R_1 > R_{\text{unequal}}$.

iv. If $\theta_H < \theta < \theta_{\text{critical}}$ and $H < T$, then $R_2 < R_{\text{unequal}}$.[21]

v. If $\theta < \theta_L$, then $R_1 < R_{\text{equal}}$.

vi. If $\theta < \theta_L$, then $R_2 > R_{\text{equal}}$.[22]

vii. If $\theta_L < \theta < \theta_{\text{critical}}$ and $L > T$, then $R_1 > R_{\text{equal}}$.[23]

viii. If $\theta_L < \theta < \theta_{\text{critical}}$ and $L < T$, then $R_2 < R_{\text{equal}}$.

Proof. Re-expressing $\theta < \theta_H$ yields

$$1 - \frac{4}{H(1-\theta)} + \frac{4}{H^2(1-\theta)^2} < 1 - \frac{4\theta}{(1-\theta)^2}. \tag{18}$$

Since both sides of this inequality are positive, taking the square root leads to the statements in (i) and (ii).

If $\theta_H < \theta < \theta_{\text{critical}}$, " $>$ " replaces " $<$ " in (17). The right-hand side is the term under the square root for R_1 and is positive since $\theta < \theta_{\text{critical}}$ (recall the proof of Lemma 2). Taking the square root implies

$$\left|1 - \frac{1}{H}\frac{2}{1-\theta}\right| > \sqrt{1 - \frac{4\theta}{(1-\theta)^2}}. \tag{19}$$

If $H > T$, then $1 - 2/[H(1-\theta)]$ is positive, and (iii) follows. If $H < T$, then $1–2/[H(1-\theta)]$ is negative, and (iv) follows. The proofs of (v)–(viii) are similar to those of (i)–(iv), respectively, where L replaces H. ■

Proving the relationships in Figure 3 is a straightforward application of Propositions 1 and 2 and Lemmas 1–3:

Proposition 4

(Column A) *If θ is less than both θ_H and θ_L, only R_{unequal} is admissible.*
(Column D) *If $\theta > \theta_{\text{critical}}$, only R_{equal} is admissible.*
(1B and 2B) *If $\theta_L < \theta < \theta_H$, then only R_{equal}, R_2, and R_{unequal} are admissible.*
(3B and 4B) *If $\theta_H < \theta < \theta_L$, then only R_1 is admissible.*
(1C) *If $L < H < T$ and $\theta_H < \theta < \theta_{\text{critical}}$, then only R_{equal} is admissible.*
(2C and 3C) *If $L < T < H$ and $\theta_L < \theta_H < \theta$, then only R_{equal}, R_2, and R_1 are admissible. Or, if $L < T < H$ and $\theta_H < \theta_L < \theta$, then only R_{equal}, R_2, and R_1 are admissible.*
(4C) *If $T < L < H$ and $\theta_L < \theta < \theta_{\text{critical}}$, then only R_{equal} is admissible.*

Proof.
(Column A) Proposition 1 implies R_{unequal} is admissible. Proposition 2 implies R_{equal} is inadmissible. Lemma 3(i) implies R_1 is inadmissible. Lemma 3(vi) implies R_2 is inadmissible.
(1B and 2B) Propositions 1 and 2 imply R_{equal} and R_{unequal} are admissible. Lemma 3(i) implies R_1 is inadmissible. Lemma 4 implies $R_2 > R_{\text{unequal}}$. Lemma 2 implies $L < T$. Lemma 3(viii) implies $R_2 < R_{\text{equal}}$. Hence $R_{\text{unequal}} < R_2 < R_{\text{equal}}$ so R_2 is admissible.

The remainder of the proof is similar. ■

Note that the relationships at the cut-off values along the *X*-and *Y*-axes in Figure 3 follow immediately from the fact that Lemma 3 can also be proved with non-strict inequality.

Theorem *Assumptions A1–A3 are necessary and sufficient conditions for multiple steady states.*

- *A1:* $0 < \theta < 3 - 2\sqrt{2} \equiv \theta_{\text{critical}}$.
- *A2:* *The time cost* c^L *is sufficiently low that* $R_{\text{equal}} \in (R_2, 1)$.
- *A3:* *The time cost* c^H *is sufficiently high that* $R_{\text{unequal}} \in (0, R_2)$.

R_2 *(and* R_1*) are the positive (and negative) roots of (6) with* $\gamma_t = \theta$*, where*

$$\begin{aligned} R_1 &= \left[1 - \theta - \sqrt{1 - 6\theta + \theta^2}\right]/2, \\ R_2 &= \left[1 - \theta + \sqrt{1 - 6\theta + \theta^2}\right]/2. \end{aligned} \tag{20}$$

Proof. Assumption A1 is the equivalent of ruling out Column D. Assumption A2 is the equivalent of ruling out Columns A, 3B, 4B, and 4C. Assumption A3 is the equivalent of ruling out 1C. ■

B. Stability

When R_{equal}, R_{unequal}, and R_1 are admissible, they are generically stable as well. When R_2 is admissible, it is generically unstable. For certain proportions of children of unskilled with low cost of education, some of the steady states become saddle points; but this occurs with measure zero.

Proposition 5 *If* R_1 *and* R_2 *are admissible then they are generically stable and unstable, respectively. When* $\theta = \theta_{\text{critical}}$, R_1 *and* R_2 *are saddle points.*

Proof. Taking the derivative of (4) and substituting $R_t = R_2$ and $\gamma^* = \theta$ yields

$$\frac{\delta R_{t+1}}{\delta R_t} = 1 + \sqrt{1 - \frac{4\theta}{(1-\theta)^2}} \geq 1, \tag{21}$$

which means R_2 is unstable or a saddle point. R_2 is a saddle point when the term under the square root is zero. Since this is the discriminant of the quadratic equation (11), this is the same as saying $\theta = \theta_{\text{critical}}$ (recall the proof of Lemma 1), which occurs with measure zero.

Substituting R_1 instead of R_2 leads to the same expression as (19), except "+" becomes "−" and "≥" becomes "≤". Thus R_1 is generically a stable steady-state and a saddle point only when $\theta = \theta_{\text{critical}}$. ■

Proposition 6 *If R_{equal} and R_{unequal} are admissible then they are generically stable as well. If $\theta = \theta_L$ and $L < T$, then R_{equal} is a saddle point. If $\theta = \theta_H$ and $H < T$, then R_{unequal} is a saddle point.*

Proof. Consider R_{equal}. Suppose the ratio of skilled workers to unskilled workers is perturbed to $R_{\text{equal}} + \eta$ at time t. At time $t+1$, the ratio of skilled to unskilled must fall or else the wage differential is less than L, which is inconsistent with rational expectations since no one would have become educated in the first place.

Suppose the ratio of skilled to unskilled is perturbed to $R_{\text{equal}} - \eta$ at time t. Assume towards contradiction that the ratio of skilled to unskilled does not rise at time $t+1$, in which case the wage differential will be greater than L. Then all θ children of unskilled workers with low cost to education would have become educated. But for R_{equal} to be admissible, $\theta \geq \gamma_{\text{equal}}$, which means at least as many children become skilled at $R_{\text{equal}} - \eta$ as at R_{equal}. For η small enough, if more people are becoming educated at $R_{\text{equal}} - \eta$ than at R_{equal}, this means the ratio of skilled to unskilled workers actually increases at $R_{\text{equal}} - \eta$, which is our contradiction.

If, however, the same proportion of people are becoming educated at $R_{\text{equal}} - \eta$ as at R_{equal} and $L < T$, then the ratio of skilled to unskilled falls at $R_{\text{equal}} - \eta$. $L < T$ and the fact that $\gamma_{\text{equal}} = \theta_L$ imply $R_{\text{equal}} = R_2$ (using Lemma 3(vi) and (viii) with non-strict inequality), which is generically unstable. Hence R_{equal} is a saddle point. If, however, $\theta = \gamma_{\text{equal}}$ and $L \geq T$, then $R_{\text{equal}} = R_i$, which is also stable. The bottom line is that R_{equal} is stable or a saddle point that occurs with measure zero.

Note that we can rule out exotic dynamics such as cycles or chaotic dynamics. For ratios of skilled workers to unskilled workers that are greater than R_{equal}, the slope of the function $R_{t+1}(R_t)$ is necessarily zero because of rational expectations. For ratios that are less than and arbitrarily close to R_{equal}, γ is set under rational expectations so that $R_{t+1} = R_{\text{equal}}$. For ratios that are sufficiently less than R_{equal}, γ is a constant and taking the derivative of (4) yields an expression that is positive.

The proof for R_{unequal} is similar. ■

Note that the comparative statics for the cutoff values, T, θ_L, θ_H, and θ_{critical}, in Figure 3 are less straightforward. For low values of L, increases in L can eliminate the equal steady state if increasing L causes θ_L to increase beyond θ. The intuition behind this is that higher L means that it becomes more difficult for people to become skilled and thus it becomes more difficult to maintain the equal steady state. Analogously, reductions in H can eliminate the unequal steady states, R_{unequal} and R_1, if H is reduced to below T and θ_H falls below θ. Lower H means it becomes easier for some people to become skilled and thus unequal steady states become more difficult to maintain.[24]

Acknowledgments

We are grateful to George Akerlof, Yu-chin Chen, Kevin Fletcher, Seema Jayachandran, Larry Katz, Paras Mehta, Petia Topalova, David Weil, and a very helpful anonymous referee for many suggestions. We have also benefitted from suggestions by seminar participants at Chicago, Michigan, MIT, and the Federal Reserve. Michael Kremer thanks the MacArthur Foundation and the National Science Foundation for financial support. Daniel L. Chen also thanks the National Science Foundation for financial support. Views expressed in this paper should not be attributed to these institutions.

Notes

1. Note that the Cobb–Douglas production form implicitly assumes that the elasticity of substitution between skilled and unskilled workers is 1. The literature suggests that the elasticity of substitution may be close to 1; see Krusell et al. (1997) and the references within for further detail. For expositional clarity, factor shares are assumed to be $\frac{1}{2}$.
2. An alternative explanation for the negative relationship between education and fertility is that parents substitute quality of children for quantity. Educated parents may have a lower shadow price of child quality, for instance, through easier access to schooling and health care; this raises demand for child quality and reduces demand for child quantity. While this may also play a role, there is empirical evidence for the view that substitution effects are important. Since women usually bear most of the responsibility for childcare, the opportunity cost of having children for women is higher than the opportunity cost of having children for men. Assuming that education is a satisfactory proxy for lifetime wage rates, this implies that the fertility differential between educated and uneducated women should be higher than the fertility differential between educated and uneducated men. Schultz (1981) and Razzaz (1998) both cite evidence confirming that female education has a strong negative effect on fertility while male education has a smaller, less statistically significant, and sometimes positive effect on fertility (United Nations, 1987 also has corroborating evidence).
3. One way to derive fertility from preferences is to assume the quasilinear utility function $V = \ln\,(n) + X$, where n is the number of children and X is consumption. Raising each child requires a time commitment of ϕ, and the total time endowment of each individual is 1. Thus the budget constraint is $X = w\,(1 - n\phi)$. Substituting X into the utility function yields $V = \ln\,(n) + w\,(1 - n\phi)$. The first order condition for optimal fertility results in equation (2). Under the assumed quasi-linear utility function, higher wages lead people to have fewer children. Note that if the utility function were $V = \ln\,(n - \varepsilon) + X$, instead of the number of children asymptoting to 0, the number of children asymptotes to $\varepsilon > 0$. This is one possible micro-foundation for the discussion in Section 4.

   A weakness of this formulation is that for tractability we derive wages under the assumption that individuals supply their entire non-schooling time to the labor market, whereas this micro-foundation explicitly considers time spent raising children. Solving the dynamic model becomes much more difficult if relative wages depend on the skilled-to-unskilled ratio of market labor time, not skilled-to-unskilled ratio of the population. R_t would depend not only on R_{t-1} but also on R_{t+1}, which takes into account fertility rates, and thus time spent in the labor market, in period t. We conjecture that most of the intuition would go through in such a model.
4. Note that ideally we would be looking at differential number of surviving infants between educated and uneducated women; however, data for differential infant mortality is scarce and in any case, differential infant mortality is small relative to differential fertility (see Section 4).
5. If children of skilled parents require $\delta < c^{\mathrm{L}}$ units of time to obtain education, the steady state proportions of skilled workers and the steady state levels of inequality are identical. If a proportion of children of skilled parents also require c^H units of time to become skilled, but the proportion is less than $1 - \theta$, the proportion of children of unskilled who require c^H units of time, the basic story is similar but (6) becomes a cubic equation instead of a quadratic equation and we no longer get closed form solutions such as those of Proposition 1.

6. This proof implies that the steady states are not Pareto rankable.
7. We did not try to look directly at the impact of ratio of educated to uneducated people on fertility differentials because of the difficulty in defining ''educated'' people cross-nationally.
8. For an in-depth discussion of how the data sets were put together, see Chen (1999).
9. For United Nations (1987) and (1995), length of attendance to formal schooling was grouped into four categories: 0 years, 1–3 years, 4–6 years, and 7 years or over. Women with 1–3 years of schooling have typically attended but not completed the primary level, those with 4–6 years of schooling have usually completed a significant portion of the primary level, and those with 7 or more years of schooling are likely to have progressed to secondary school. Collapsing years of schooling into a few categories tends to reduce the response error. For certain data, the educational group 7 years or over is divided into two categories, 7–9 years and 10 years or over, because of the recent upward trend in female education and thus to provide more detail at the upper end of the educational spectrum. For Mboup and Saha (1998) and Jones (1982), the education categories reported are descriptive, for example, no school, primary, and secondary+.
10. Note that the percentage distribution of women is only available for data from United Nations (1987, 1995).
11. In an attempt to use data as consistent as possible across countries and across time, we avoid using individual country censuses, which presumably would vary significantly in definitions and measurement. To reduce measurement error, we use data sources, such as the World Fertility Survey and Demographic and Health Surveys, that have already standardized definitions and measurements across many countries.
12. This factor is calculated by computing $e^{(10^{*}\text{fertility differential increase})}$.
13. We also find a positive relationship between differential fertility and Mincer coefficients of returns to education in a data set of 30 countries for which both variables could be obtained. This relationship is less significant perhaps due to the small sample size and noise in computing Mincer coefficients.

    We also find a positive relationship between differential fertility and income inequality in a United States time-series data set consisting of 13 observations for the United States from 1925 to 1989. That the relationship is less significant than in the other data sets can perhaps be explained by the small size of the United States time-series sample. Furthermore, as Section 4 notes, the relationship may be hard to pick up in general for a high-income country such as the United States. Since fertility levels are generally very low, measurement noise may cover up the true association between differential fertility and income inequality.
14. To compare with Table 2, the weighted OLS coefficient measuring differential fertility is -0.107.
15. See footnote 4 for one possible micro-foundation.
16. Note that \$1,639 is the World Bank cutoff between lower-middle and upper-middle income countries.
17. The empirical investigation of the model can also be expanded but evidence is limited by the available data.
18. The model is such that the equilibrium ''snaps'' into place in one step if it is in the neighborhood of a stable steady state. Hence, the slope of the response function $R_{t+1}(R_t)$ in the neighborhood of the steady state is zero. The traditional force of children choosing education when wage differentials are high has an infinite weight relative to the demographic force and hence the supply of skilled labor is infinitely elastic.
19. Due to space limitations, intermediate steps in many of the subsequent proofs are also omitted; the reader can refer to Chen (1999) for more detail.
20. An equivalent necessary and sufficient condition for admissibility of R_1 and R_2 in Proposition 3 is $R_{\text{unequal}} \le R_1$ and $R_2 \le R_{\text{equal}}$. This is because $w_t^u/w_t^s = R_t$ and because the expressions in Propositions 1 and 2 imply that the wage differentials at R_1 and R_2 between L and $H \Leftrightarrow R_1$ and R_2 between R_{unequal} and R_{equal}. In other words, for a steady state to be admissible, the skilled-to-unskilled ratio must be between the skilled-to-unskilled ratios at the equal and unequal steady states. Thus Lemma 3 can be interpreted as saying if R_{unequal} is admissible, then R_1 is inadmissible.
21. Hence, neither R_2 nor R_1 is admissible.
22. In other words, if R_{equal} is inadmissible, then so is R_2.
23. Hence neither R_2 nor R_1 is admissible.
24. For large values of L, however, increases in L can eliminate R_1 and expand the region of θ over which only the equal steady state is admissible, if increasing L causes θ_L to decrease below θ. This is because $\theta_L(L)$ is increasing in L until $L = T$ and then decreasing thereafter (Lemma 2). This may seem counter-intuitive. However, recall that increasing L and H increases the wage differential at the equal and unequal steady states, respectively. For example, it may seem odd that increasing L from 3C to 4C in Figure 3 removes the unequal steady states. Or it may seem odd that increasing L so that θ_L falls below θ may move the system from 4B to 4C, from R_1 to R^*_{equal}. But for R_1 to be admissible in the first place, its wage differential must have been

between L and H. Thus, increasing L increases the lower bound of admissibility for R_1. When increases in L cause R_1 to become inadmissible, the economy then shifts to R^*_{equal} but the wage differential, L, at R^*_{equal} is now greater than it was at R_1. A similar phenomenon occurs when increases in H cause the system to shift from 4A to 4B by lowering θ_H below θ.

References

Barro, R., and X. Sala-i-Martin. (1995). *Economic Growth*. New York: McGraw-Hill.

Birdsall, N. (1988). ''Economic Approaches to Population Growth,'' in Chenery, H., and Srinivasan, T.N., eds., *Handbook of Development Economics*, vol. 1. Amsterdam: Elsevier Science Publishers, pp. 477–542.

Chen, D. L. (1999). ''Income Distribution Dynamics with Endogenous Fertility,'' Senior Thesis, Harvard University.

Chu, C. Y. C., and H.-W. Koo. (1990). ''Intergenerational Income-Group Mobility and Differential Fertility,'' *American Economic Review* 80(5), 1125–1138.

Dahan, M., and D. Tsiddon. (1998). ''Demographic Transition, Income Distribution, and Economic Growth,'' *Journal of Economic Growth* 3(1), 29–52.

Deininger, K., and L. Squire. (1996). ''Measuring Income Inequality: A New Data-base.'' Development Discussion Paper No. 537, Harvard Institute for International Development, Harvard University.

Galor, O., and J. Zeira. (1993). ''Income Distribution and Macroeconomics,'' *Review of Economic Studies* 60(1), 35–52.

Galor, O., and H. Zang. (1997). ''Fertility, Income Distribution, and Economic Growth: Theory and Cross-Country Evidence,'' *Japan and the World Economy* 9(2), 197–229.

Jones, E. (1982). *Socio-economic Differentials in Achieved Fertility.* World Fertility Survey Comparative Study No. 21, ECE Analyses. Voorburg, The Netherlands: International Statistical Institute.

Kremer, M., and D. Chen. (2000). ''Income Distribution Dynamics with Endogenous Fertility,'' NBER Working Paper No. 7530.

Krusell, P., L. Ohanian, V. Rios-Rull, and G. Violante. (1997). Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis. Federal Reserve Bank of Minnesota Staff Report #239.

Lam, D. (1986). ''The Dynamics of Population Growth, Differential Fertility, and Inequality,'' *American Economic Review* 76(5), 1103–1116.

Mare, R. D. (1997). ''Differential Fertility, Intergenerational Educational Mobility, and Racial Inequality,'' *Social Science Research* 26(3), 263–291.

Mboup, G., and T. Saha. (1998). *Fertility Levels, Trends, and Differentials: Demographic and Health Surveys*, Comparative Study No. 28. Calverton, Maryland: Macro International.

Moav, O. (2001). ''Cheap Children and the Persistence of Poverty.'' Mimeo.

Morand, O. (1999). ''Endogenous Fertility, Income Distribution, and Growth,'' *Journal of Economic Growth* 4(3), 331–349.

Nei cheng pu. (1996). *Chung-hua min kuo Tai Min ti chu jen kou tung chi*. Taiwan-Fukien demographic fact book, Republic of China, Tai-pei.

Nerlove, M., A. Razin, and E. Sadka. (1984). ''Income Distribution Policies with Endogenous Fertility,'' *Journal of Public Economics* 24(2), 221–230.

Preston, S., and C. Campbell. (1993). ''Differential Fertility and the Distribution of Traits: the Case of I.Q,'' *American Journal of Sociology* 98(5), 997–1019.

Razzaz, S. (1998). ''Fertility and Growth: The Significance of Gender Differences in Human Capital.'' Unpublished Ph.D. dissertation, Brown University.

Repetto, R. (1978). ''Interaction of Fertility and Size Distribution of Income,'' *Journal of Development Studies* 14(4), 22–39.

Schultz, T. P. (1981). *Economics of Population*. Massachusetts: Addison-Wesley.

Summers, R., and A. Heston. (1991). ''The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988,'' *Quarterly Journal of Economics* 106(2).

United Nations. (1987). *Fertility Behavior in the Context of Development: Evidence from the World Fertility Survey.* Population Studies No. 100, New York: United Nations, 1987.

United Nations. (1995). *Women's Education and Fertility Behavior: Recent Evidence From the Demographic and Health Surveys*. New York: United Nations.

World Bank. (1999). *World Development Report 1999/2000*. Oxford: Oxford University Press.