A Theory of Occupational Choice with Endogenous Fertility

By Dilip Mookherjee, Silvia Prina and Debraj Ray*

Theories based on partial equilibrium reasoning alone cannot explain the widespread negative cross-sectional correlation between parental wages and fertility, without restrictive assumptions on preferences and childcare costs. We argue that incorporating a dynamic general equilibrium analysis of returns to human capital can help explain observed empirical patterns. Other by-products of this theory include explanations for intergenerational mobility without stochastic shocks, connections between mobility and fertility patterns, and locally determinate steady states. Comparative statics exercises on steady states shed light on the effects of education, childcare subsidies, child labor regulations and income redistribution policy on long run living standards.
JEL: J13, J24, J62
Keywords: Inequality, human capital, endogenous fertility, borrowing constraints.

A widespread empirical finding in the economics of fertility and human capital is a negative cross-sectional relationship between parental wages and fertility.[1] Wealthier parents tend to have fewer children who are better educated, suggesting parents tend to substitute quality for quantity of children as their wages increase. The final phase of the demographic transition traces a similar phenomenon over time, as fertility rates fall with rising per-capita incomes. Beginning with the work of Becker (1960), a large theoretical literature attempts to explain these findings, by modeling the decisions of parents who trade off the benefits of having more children against the attendant costs of childcare, including the reduction in time available for work.

However, with few exceptions, this entire literature imposes restrictive assumptions on preferences that restrict the strength of wealth effects relative to the substitution effects associated with increases in parental wages, in order to explain a negative effect on fertility.[2] Typically, utility functions over parental

* Mookherjee: Department of Economics, Boston University, 270 Bay State Road Boston, MA 02215 (email: dilipm@bu.edu); Prina: Department of Economics, Case Western Reserve University, 11119 Bellflower Road, Cleveland, OH 44106 (email: silvia.prina@case.edu); Ray: New York University, Department of Economics, 19 West 4th Street New York, NY 10012 (debraj.ray@nyu.edu). Mookherjee and Ray acknowledge funding from the National Science Foundation under grant nos. SES-0617874 and SES-0617827 respectively. Ray is additionally grateful for funding from the Fulbright Foundation and for hospitality from the Indian Statistical Institute during a year of leave from NYU.

[1] For an excellent overview, see Jones, Schoonbroodt and Tertilt (2008).

[2] One interesting exception is an approach in which parental wealth — via human capital investments — and fertility are determined simultaneously (Jones, Schoonbroodt and Tertilt 2008). Suppose there

consumption must exhibit smaller curvature than a logarithmic function; alternately the elasticity of substitution between consumption and fertility must be less than one. When the theory is extended to incorporate investments in child quality (as in Becker and Tomes 1976; Moav 2005) this approach generates a negative wage-fertility relationship only if restrictive assumptions on childcare costs are additionally imposed.[3]

It should be noted that the negative relationship between wages and fertility though widespread, is not universal. In developing countries, or in the historical record of currently developed countries, a positive relationship cannot be discarded altogether.[4] But it is not just a question of variation; it is variation of a systematic kind. In the context of cross-sectional US data, Freedman (1963) found that *within the same occupation*, higher income households tend to have more children, while overall — not controlling for occupation — average fertility varies negatively with income.[5] This is consistent with observations made earlier by Spengler (1952) and Easterlin (1973) that fertility growth is positively related to relative incomes within an occupational category. Simon (1969) shows similar evidence using 1964 US household data showing that fertility rates varied positively with incomes within most occupation-region combinations. Simon also pointed out that over the course of the business cycle the correlation of fertility with income tends to be positive, in contrast to the cross-sectional pattern, also possibly for these reasons.[6] It seems difficult to explaining these exceptions to the general pattern in terms of different relative strengths of wealth and substitution effects; there is no particular reason for these to differ between inter-occupation and intra-occupation comparisons.

Indeed, Becker (1960) himself claimed that the wealth effect was typically dominant, based on an examination of the US evidence available at that time. He argued that the negative cross-sectional relationship was an artifact, caused by ignorance of contraceptive methods on the part of low-income individuals.

is unobserved heterogeneity in parental preferences for fertility, and fertility decisions affect the human capital investments of parents. Then parents who have more children will have lower human capital and earn lower wages. Causality now effectively runs from fertility to parental wages. This explanation is less satisfactory in the context of developing countries, where there is limited scope for human capital investments by adults, and human capital attainments are largely determined by early-childhood investments made by one's parents.

[3]The goods-cost of childcare must be small relative to the innate abilities of children, implying a negative goods cost per child, net of variable schooling costs. See Jones, Schoonbroodt and Tertilt (2008, section 5.2), and a more detailed explanation in Section V below.

[4]These include Simon (1977) for Poland in 1948; Clark (2005), Clark and Hamilton (2006), and Clark (2007) for England in the 16th and 17th century; Weir (1995) for France in the 18th century; Wrigley (1961) and Haines (1976) for some areas in France and Prussia in the 19th century; and Lee (1987) for the U.S. and Canada.

[5]This paper used a 1955 US nationally representative sample of married non-farm women between the ages of 18 and 39, and found a positive relationship between fertility and income of the husband relative to average income of others similar in age, education and occupation. On the other hand, fertility and husband income were negatively related.

[6]More recently, Blau and van der Klaauw (2007) describe a negative effect of wages on fertility for white male wage earners, but a positive effect for Blacks and Hispanics, in the 1979 cohort of the National Longitudinal Survey of Youth.

In this paper we take a somewhat different approach, in which occupational shifts play an essential role in determining the wage-fertility correlation. We study endogenous fertility in a model of occupational choice in which human capital investments are made by parents when agents are in their childhood. Parents choose fertility as well as human capital for their children, motivated by standard non-paternalistic parental altruism *a la* Barro and Becker (1986, 1988, 1989). In this respect we follow the lead of Doepke (2004). In contrast to Doepke (2004), however, we impose no restrictions on preferences concerning the relative magnitude of wealth and substitution effects, or on the nature of childcare or education costs. Instead, we analyze the steady states of a dynamic general equilibrium model in which the returns to human capital are endogenously determined. Our central finding is that the resulting general equilibrium model can help explain observed empirical patterns when there are multiple occupations with varying human capital requirements. Specifically, *across occupations the correlation is negative in general, while within occupations it is positive if wealth effects are strong relative to substitution effects.* The model therefore explains the widespread negative correlation between wages and fertility, while allowing for exceptions that arise when we look within occupations, or in contexts with limited scope for human capital investments.

It is important to be explicit about the precise role that general equilibrium plays in our results. In partial equilibrium, parental wages are exogenously given, and an increase in that wage — say, as we move over the occupational cross-section — has two kinds of effects on fertility. One is an *occupational shift* effect: rising wages induce parents to invest more in education of their children, which *always* tends to reduce fertility.[7] The other is the familiar pure *preference* effect which operates if we control for the education decision: with unchanged education, rising wages induce a rise in fertility if wealth effects are strong relative to substitution effects. With strong wealth effects, then, the net outcome is still indeterminate in partial equilibrium, though it is certainly tilted more strongly in favor of declining fertility relative to a model with no occupational choice. In particular, much depends on the size of the skill premium: the magnitude of wage differences between educated and uneducated parents.

In general equilibrium, the skill premium is endogenously determined. If wealth effects are strong and fertility is rising across occupations, the steady-state skill premium must be low enough to ensure that the greater supply of children from high-wage parents is suitably spread out across occupations. We show that this restricts the skill premium sufficiently, so that the preference effect is dominated by the occupational shift effect. The same result obtains when wealth effects are small, or when the goods component of child-care costs are negligible. In summary, the general equilibrium approach is fundamental to our exercise because these endogenous restrictions allow us to "sign" the cross-sectional relationship between parental wages, fertility and skill choice in an unambiguous way.

[7]This result also appears in Doepke (2004).

Our model fuses endogenous fertility theory with standard dynamic general equilibrium models of occupational choice with borrowing constraints, which assume constant and exogenous fertility.[8] We find that incorporating endogenous fertility helps resolve a number of problems with the standard model. One is that intergenerational mobility arises endogenously in steady state despite the absence of any stochastic shocks. Existing theories of mobility that assume constant, exogenous fertility (e.g. Becker and Tomes 1979; Loury 1981; Banerjee and Newman 1993; Mookherjee and Napel 2007) rely on stochastic shocks to abilities or incomes. But when fertility differs across occupations, we show that mobility is necessary to ensure the constancy of steady state skill ratios across successive generations. For instance, if fertility is higher in unskilled occupations, the proportion of skilled agents in the economy will tend to drift downwards over time. A steady state in which per capita skill in the economy is constant over time therefore requires upward mobility: a fraction of unskilled households must decide to educate their children to prepare them for entry into the skilled occupation. A society with a higher fertility differential between the skilled and unskilled must therefore involve greater mobility. This connection between mobility and fertility patterns has been overlooked in the existing literature on intergenerational mobility.[9]

Endogenous fertility also generates local determinacy of steady states, in contrast to occupational choice models with exogenous fertility. By the logic sketched above, steady states with differential fertility across occupations must be associated with the indifference of parents in one occupational category between educating and not educating their children. This indifference condition ties down relative wages and the skill ratio in steady state, ensuring the local determinacy of macroeconomic aggregates, thus permitting a tractable analysis of the long-run effects when key parameters change in the model.[10] Our theory generates predictions about the macroeconomic effects of childcare or education subsidies, redistributive tax-transfer policies or child labor regulations. Specifically, a rise in the non-time component of childcare costs, or a fall in education costs, or stronger child labor regulations, are shown to increase long run human capital investments, raise per-capita income and lower wage inequality across skilled and unskilled occupations. The same effects obtain with a *reduction* in unconditional transfers to the unskilled that are funded by taxes on earnings of the skilled, or an *increase* in transfers conditioned on school enrollment of children. These results suggest that some key aspects of the demographic transition (such as tendency for fertility to fall with urbanization (which raises childcare costs, lowers education costs),

[8]Examples of the latter include Banerjee and Newman (1993), Galor and Zeira (1993), Ljungqvist (1993), Freeman (1996), Aghion and Bolton (1997), Lloyd-Ellis and Bernhardt (2000), Matsuyama (2000, 2003), Ghatak and Jiang (2002) and Mookherjee and Ray (2002, 2003, 2010).

[9]See, for example the symposium in the *Journal of Economic Perspectives* 2002, 16(3).

[10]In this respect the analysis also contrasts with the representative-agent growth model of Barro and Becker (1989) with a perfect capital market, in which long run policies tend not to have any long run effects.

increasing child labor regulations or access to family planning) can be explained by comparative static properties of steady states, rather than non-steady-state dynamics which forms the standard approach in the literature.[11]

The paper is organized as follows. Section I introduces the model. Section II analyzes household optimal choices in the partial equilibrium setting with given wages and continuation values of children. Following this, Section III introduces competitive equilibrium, characterizes properties of steady states, including the relation between fertility and mobility patterns, and provides conditions under which a negative wage-fertility correlation obtains. Section IV shows steady states are locally determinate and performs comparative static exercises. Section V then discusses relation to existing literature in some detail. Finally Section VI concludes, while the Appendix collects proofs.

I. Model

A single output is produced under competitive conditions, using skilled and unskilled labor. Let λ denote the fraction of skilled labor. The marginal product of skilled labor decreases in λ; the opposite is true of unskilled labor. Both marginal products are smooth functions of λ and satisfy Inada endpoint conditions.[12]

There are two occupations, unskilled (0) and skilled (1). Skilled workers can work as either skilled or unskilled labor, a choice not available to unskilled workers. Let $\bar{\lambda}$ denote the value of λ for which the marginal products of skilled and unskilled labor are equalized. Then for $\lambda < \bar{\lambda}$, skilled and unskilled wages (w_1 and w_0) equal their respective marginal products, while for higher skill ratios they both equal the common marginal product at $\bar{\lambda}$.

There is a continuum of households at every date, with one adult (a single parent) in each household.[13] A parent earns a wage w on the labor market and chooses how many children n to have, where we suppose n to be a continuous variable.[14]

Child-rearing and education are costly activities. We distinguish between the cost incurred in raising an unskilled child — $r_0(w)$ — and the cost of raising a skilled child — $r_1(w)$. We maintain the following assumptions on r_0 and r_1 throughout the paper:

[R.1]. For each category $i = 0, 1$, $r_i(w)$ is smooth and strictly increasing in w, while $\frac{r_i(w)}{w}$ is nonincreasing. There is a positive lower bound $\underline{r}$ to the rate of

[11]See papers by Dahan and Tsiddon (1998), Kremer and Chen (1999, 2002), de la Croix and Doepke (2003) and Doepke (2004) which study how following some shocks an economy may shift from one steady state to another.

[12]That is, they go to ∞ and 0 at either end of their variation.

[13]It is possible with no great gain in insight to extend the model to two parents per household.

[14]Conceivably, similar results can be obtained in a model with integer-valued family size and cross-household heterogeneity in parental fertility preferences, where we can interpret the n obtained in the current theory as the average number of children (conditional on parental economic status) that would arise in the richer model. Whether and when such "purification" can be achieved is a question for future research.

increase of $r_i, i = 0, 1$.

[R.2]. For every w, $r_1(w) > r_0(w)$.

[R.3]. For every w,

$$\frac{r_1'(w)}{r_1(w)} < \frac{r_0'(w)}{r_0(w)}.$$

Assumption R.1 states that higher parental wages increase child-rearing costs, but less than proportionately. The former arises from the time component to child-rearing which causes the parent to be away from work.[15] Other fixed resource costs of child-bearing and rearing would be independent of parental wages, which would imply that per-child costs would rise less than proportionately with parental wages. Note that R.1 allows such fixed costs to be zero.[16] Assumption R.2 is self-evident: imparting skills to children is costly. Assumption R.3 states that the *marginal* cost impact of a higher parental wealth (relative to the overall upbringing budget) is lower for skilled children than for unskilled children.

Suppose, for instance, that there is a child-rearing component $k(w)$, and an additional cost $s(w)$ of imparting skills ("k" for kids, "s" for skills). Then $r_0(w) = k(w)$ and $r_1(w) = k(w) + s(w)$, and R.1–R.3 are met provided that k and s are increasing and $k(w)/s(w)$ increases in w. In particular, R.1–R.3 hold if $s(w)$ equals some fixed constant, which can be interpreted as schooling cost. Call this cost structure *separable*.

An important subcase of the separable structure is one in which rearing each child involves only parental time, and no material goods cost. Then $k(w) = \psi w$ for some $\psi > 0$ which represents the fraction of time spent away from work in raising children. We will use this for one of the main results.[17]

While the functions r_0 and r_1 are exogenous to the model, they can be influenced by policies pertaining to child-care subsidies, child labor regulations and costs of family planning. Section 5 of the paper will examine these effects.

These assumptions generalize most formulations of childcare costs in the literature. For instance, most of the models surveyed by Jones, Schoonbroodt and Tertilt (2008) entail a linear formulation of the functions r_0 and r_1, all of which satisfy R1–R3.

Unlike most preceding models of fertility, we allow parents to educate some children but not others. Let e be the fraction of children made skilled; then the expenditure per child is given by $r(w, e) \equiv er_1(w) + (1 - e)r_0(w)$, and total expenditure on children is given by $r(w, e)n$, so that the lifetime consumption of the parent is equal to

$$c = w - r(w, e)n,$$

[15]This formulation is compatible with the possibility that skilled parents find it easier to educate their children, but we assume that the *net* monetary cost still rises with the wage.

[16]Our result concerning existence of interior steady states, however, does require the assumption of positive fixed costs.

[17]Another relevant subcase is one where $k(w) = f + \psi w$, the sum of a fixed goods cost and maternal time costs. We focus on this case in the section on comparative statics.

which we constrain throughout to be nonnegative. This reflects the underlying credit constraint common to all occupational choice models, in which education or child-rearing costs cannot be financed by borrowing and must entail consumption sacrifices by parents.

We now describe preferences. Each parent possesses, first, a utility indicator defined on lifetime consumption c, given by $u(c)$. The parent also derives utility from the lifetime payoff V to be enjoyed by each child. These latter values will be endogenous to the model and will be solved for in the general equilibrium analysis, but will be taken as given in the partial equlibrium. We write the overall payoff to a parent as

$$u(c) + \delta n^{\theta} \left[eV_1 + (1-e)V_0 \right]. \tag{1}$$

where δ is the cross-generational discount factor, n^{θ} is a weighting factor that depends on the total number of children n, e is the proportion of children who are skilled, and V_j is the expected lifetime values accruing to an individual who is placed in skill category j.

As in all the literature, we presume that, *controlling for their own consumption and for child utility,* parents prefer more children to less.[18] This means that θ and the value functions (and consequently the utility function u) must have the same sign.[19] We therefore assume that

[U.1]. u is smooth, increasing, strictly concave, and has unbounded steepness when consumption is zero.

[U.2]. $\theta \neq 0$, and $\theta < 1$.

[U.3]. u is nonnegative throughout when $\theta > 0$, and negative throughout when $\theta < 0$.

[U.4]. $0 < \delta < \underline{r}^{\theta}$.

Assumption U.1 is standard. In U.2, the restriction that $\theta \neq 0$ means that parents are sensitive to family size, while the assumption that $\theta < 1$ reasonably imposes diminishing marginal returns to family size.[20] Assumption U.3 embodies the discussion that follows equation (1) above. When $\theta > 0$ and $\rho < 1$, the quantity and quality of children are complements in parental preferences. Conversely, when $\theta < 0$ and $\rho > 1$, they are substitutes. Finally, U.4 constrains the discount factor so as to ensure that value functions are well-defined. In particular, the non-negativity of parental consumption and R.1 ensure an upper bound $\frac{1}{\underline{r}}$ to fertility.[21] Hence U.4 ensures that δn^{θ} always lies between 0 and 1.

[18]We hasten to add that this does not mean that parents will *unconditionally* prefer more children to less! But it does exclude the case in which parents believe that never having been born is a better option than life.

[19]Jones and Schoonbroodt (2009) contains more discussion on the joint restrictions that link θ and u, and on the need for u to have a single sign (thereby ruling out, say, the case of logarithmic preferences).

[20]No corresponding restriction on θ needs to be imposed when it is negative.

[21]By R.1, child-rearing costs are at least $\underline{r}.wn$, so parental consumption is at most $w[1 - \underline{r}.n]$.

II. Partial Equilibrium Analysis

Fix lifetime values V_1 and V_0 for children, with $V_1 > V_0$. In a partial equilibrium context, these are taken as given, as well as the wage w of the parent, who has to decide on n, the number of children, and on e, the proportion of these children who will be educated. We proceed in two steps. First taking e as given, the parent with wage w selects fertility $n(w,e)$.[22] Then at the second step the parent selects e, incorporating the effect of e on the associated fertility decision. This is an inversion of the common-sense sequencing but makes absolutely no difference to the analysis as the parent is fully time-consistent.

At the first step, $n(w,e)$ is determined by the solution to the following first-order condition

$$u'\left(w - r(w,e)n(w,e)\right) r(w,e) = \theta n(w,e)^{\theta-1}\left[eV_1 + (1-e)V_0\right]. \tag{2}$$

This condition defines the function $n(w,e)$ uniquely at all $w > 0$.[23]

With $n(w,e)$ determined in this way, we can proceed to the second step and express parental utility as a function of e alone:

$$\begin{aligned} V(w,e) &\equiv u\left(w - r(w,e)n(w,e)\right) + n(w,e)^{\theta}\left[eV_1 + (1-e)V_0\right] \\ &= u\left(w - r(w,e)n(w,e)\right) + \frac{1}{\theta}u'\left(w - r(w,e)n(w,e)\right) r(w,e)n(w,e) \\ &= u(w-z) + \frac{1}{\theta}u'\left(w-z\right) z, \qquad (3) \end{aligned}$$

where $z \equiv r(w,e)n(w,e)$ denotes total expenditures on children, and the second equality invokes the first-order condition (2).

The second stage optimization problem involves an essential nonconvexity, owing to the form of parental preferences which depend multiplicatively on quantity and quality choices. One manifestation of this is the following observation.

LEMMA 1: *When quality and quantity of children are complements in parental preferences ($\theta > 0$), it is optimal for the parent to select e to maximize total expenditure on children, $r(w,e)n(w,e)$. Conversely, when quality and quantity are substitutes ($\theta < 0$), it is optimal for the parent to select e to minimize $r(w,e)n(w,e)$.*

To explain this, note that (3) expresses parental utility as a function of total child expenditure z alone (besides parental consumption), with a welfare weight ($\frac{1}{\theta}u'(w-z)$) that depends on the level of expenditure. In the case where $\theta > 0$,

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

[23] Obviously, $n(0,e) = 0$. Also, note that the second-order condition for this maximization problem — given e — is always met. We approach the joint determination of e and n in the main text to follow.

the welfare weight on expenditures *rises* with the level of expenditures. This is the intuition underlying the first part of the above Lemma. In the case where $\theta < 0$, the welfare weight on expenditures is negative, implying it is optimal for parents to minimize expenditures. The proof of the Lemma is straightforward: note that the expression

$$u(w-z) + \frac{u'(w-z)z}{\theta} \tag{4}$$

is strictly increasing in z under $\theta > 0$, and is strictly decreasing in z under $\theta < 0$.[24]

This property, which plays a key role in the analysis that follows, should not be misinterpreted. It *does not* state that a parent maximizes or minimizes expenditure on children by choosing e *and* n. Rather, it states that a parent maximizes or minimizes expenditure *through the choice of* e *alone*, under the artificial presumption that she "then" chooses fertility n to maximize overall payoff.

We now arrive at the main result of this section which characterizes the solution to the partial equilibrium problem. Detailed proofs of this and subsequent propositions are provided in the Appendix, while we provide the intuitive reasoning in the text following each result.

PROPOSITION 1: *(a) A parent must always set* e *equal to 0 or 1.*
(b) If there exists wage w^* *where the parent is indifferent between* $e = 0, 1$*, then* $e = 0$ *(resp.* $e = 1$*) is the unique optimum choice of* e *for all* w *below (resp. above)* w^**.*
(c) $n(w^*, 1) < n(w^*, 0)$*, i.e., fertility must drop if the parent with wage* w^* *switches from* $e = 0$ *to* $e = 1$*.*

Result (a) states that parents must decide either to educate all their children or none at all. It stems from quasi-convexity in e of the indirect utility function $V(w, e)$. At any possible interior turning point in e for parental utility $V(w, e)$, the marginal (utility) cost to parents of increasing e turns out to be *decreasing* in e. This owes in turn to a cutback in n as e rises.[25]

[24] The derivative of (4) equals

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

which is strictly positive in z under $\theta > 0$ and strictly negative in z under $\theta < 0$.

[25] Denoting the marginal (financial) cost of education by $x \equiv r_1(w) - r_0(w)$, the marginal utility cost of education for the parent is $u'.n.x$, where u' denotes the marginal utility of parental consumption, and $n \equiv n(w, e)$ denotes fertility. Using the first order condition determining the optimal fertility decision $n(w, e)$, the marginal utility cost of education can be expressed as

$$x.\theta.\delta.n^{\theta}[\frac{V_0 + e(V_1 - V_0)}{r_0 + ex}]$$

which is decreasing in e if the return to education is small in the sense that $V_1 - V_0 < V_0 \frac{x}{r_0}$. This condition is easily verified at any interior turning point for V with respect to e. It is easily verified from

Result (b) states that at any wage where the parent is indifferent between educating all or none of his children, the education option must be associated with a strictly lower fertility. This result which also appears in Doepke (2004, Proposition 2), follows from Lemma 1 above, wherein parents must either maximize or minimize expenditures on children by selecting e. Hence a shift in occupation selected for children (with wage fixed at w^*) must be associated with a discrete drop in fertility.

Result (c) states that the incentives to educate children must be (weakly) higher for parents with higher wages, a result that also appears in Doepke and Zilibotti (2005, Lemma 1). The standard argument for this result in the context of models with exogenous fertility is that the marginal utility sacrifice associated with spending (a given amount) on education must be lower for parents with higher wages, owing to borrowing constraints and the concavity of utility with respect to parental consumption. This argument does not extend straightforwardly to the current context of endogenous fertility, since the costs of education depend on parental wages. Nevertheless the result still ends up being generally true, since the incremental financial cost associated with the education option $r_1(w)n(w,1) - r_0(w)n(w,0)$ is locally decreasing with respect to w (at w^*) owing to assumption R3. In turn it ensures that there can be at most one threshold wage w^* where parents switch from one option to the other.

Combining the three parts, we infer that it will be optimal for low wage parents to not educate any of their children, while it will be optimal for high wage parents to educate all of theirs (excluding the possibility that education incentives are universal or totally absent in the population, as we will see must be the case in any competitive equilibrium since both occupations are essential in the production sector). At the wage w^* where education incentives change, a local increase in wages must entail a drop in fertility (as fertility will change continuously with respect to the wage on either side of w^*). This is what we refer to as the *occupational shift effect*, representing a tendency for quality and quantity of children to be negatively related.

Nevertheless, this is a local result, in a neighborhood of the threshold wage w^*. How does fertility vary with wages on either side of w^*, where education decisions are not changing? The answer to this is provided by studying the first order condition (2) associated with the first stage of the household optimization exercise. Here the familiar contrast between wealth and substitution effects arises. To show this contrast clearly, we restrict the nature of preferences and childcare costs.

PROPOSITION 2: *Suppose u displays constant elasticity:*

$$u(c) = \frac{c^{1-\rho}}{1-\rho},$$

the first order condition (2) that an increase in e causes u' to increase. Hence n must fall sufficiently to ensure that $u'.n$ falls.

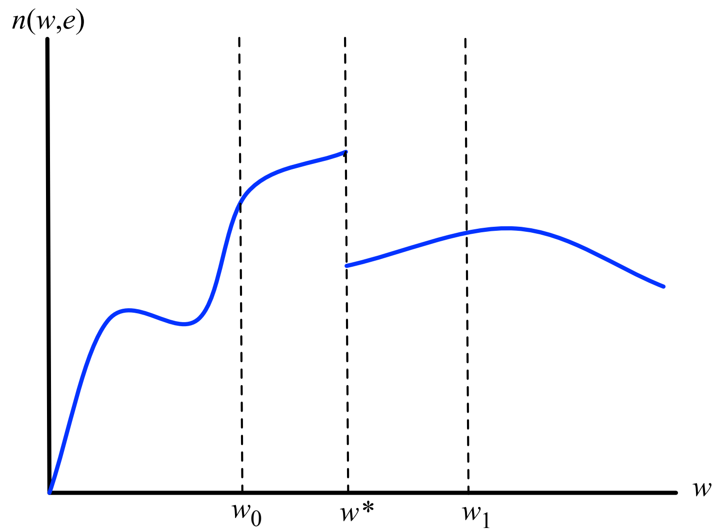

FIGURE 1. FERTILITY AT DIFFERENT PARENTAL WAGES.

where ρ is positive but not equal to 1. Presume, moreover, that for each occupational category i, rearing/educational costs are of the form

$$r_i = f_i + \psi_i w,$$

where $f_i \geq 0$ and $\psi_i \in (0,1)$. Let $\epsilon_i(w) \equiv \psi_i w/(f_i + \psi_i w)(< 1)$ denote the elasticity of r_i with respect to its argument w. Then for fixed education choice $e = i$, fertility $n_i \equiv n(w,i)$ is locally increasing in parental wage w if

$$\rho > \frac{\epsilon_i(w) - \psi_i n_i}{1 - \psi_i n_i}, \tag{5}$$

and locally decreasing if the opposite inequality holds.

An implication of this result is that *the fertility-wage correlation must be positive on either side of w^* whenever* $\rho > 1$ (since $\epsilon_i(w) \leq 1$). If the utility function exhibits at least as much curvature as the logarithmic function, wealth effects associated with wage increases dominate substitution effects. With unchanging quality, the fertility-wage correlation will be positive — this is the pure *preference effect*. Rising wages are on the other hand associated with rising quality of children — this occupational shift imparts a positive correlation. Which of the two dominates is indeterminate from partial equilibrium reasoning.

Figure 1 illustrates the relevant case where the unskilled adult wage w_0 lies below w^* while the skilled wage w_1 lies above it. The net outcome will depend on how far apart the wages of educated and uneducated adults are, which will determine the strength of the preference effect. The magnitude of the fertility drop at w^* owing to the occupational shift effect is independent of this wage gap (vide Lemma 1, which shows it just depends on w^*). Hence the net effect on the wage-fertility correlation will depend on inter-occupation wage differences, which can only be determined via general equilibrium reasoning. We turn to this in the next section.

III. General Equilibrium

A study of dynamic equilibrium with dynastic households requires three avenues of closure for the model. First, skilled and unskilled wages in every period must depend on the proportion of skilled labor in that period. Second, the continuation values that parents take as given must be identified with the maximum payoffs to their children as they grow up to be adults. Finally, given the aggregate skill ratio at date t, the skill and fertility choices made by generation t must determine the aggregate skill ratio at date $t+1$.

Formally, a *dynamic competitive equilibrium* for the economy starting with skill ratio λ_0 in generation 0 is described by the following objects, satisfying the restrictions described below:

(a) *Skilled and unskilled wages* (w_{1t}, w_{0t}) and *aggregate skill ratios* λ_t at every date, with $w_{1t} = w_1(\lambda_t)$ and $w_{0t} = w_0(\lambda_t)$ the respective marginal products; the former decreasing and the latter increasing functions, with $w_1(\bar{\lambda}) = w_0(\bar{\lambda})$ and $w_1(\lambda) > w_0(\lambda)$ for all $\lambda < \bar{\lambda}$.

(b) *Household maximization* Every parent in either skill category $i = 0, 1$ will seek to maximize

$$u(w_{it} - r_j(w_{it})n)n + \delta n^{\theta} V_{j,t+1} \tag{6}$$

by choosing fertility n and educational category $j \in \{0, 1\}$ for their children;

(c) *Continuation values* V_{it} for each category i and date t must equal the maximized value of (6);

(d) *Fertility choice* by every parent in category i, at every date, conditional on the choice of skill category j for children: $n_{it}(j)$, must be an optimal solution for n to the maximization of (6), *given* category choice j for children;

(e) *Fractions of parents* in each skill category and at each date, η_{1t} and η_{0t}, that choose the skilled category for their children, with $\eta_{it} \in (0, 1)$ only if parents in category i at date t are indifferent between the two occupational options for their children.

(f) *Evolution of aggregate skill ratios*: λ_0 is given, and

$$\lambda_{t+1} = \frac{\lambda_t \eta_{1t} n_{1t}(1) + (1-\lambda_t)\eta_{0t} n_{0t}(1)}{\lambda_t[\eta_{1t} n_{1t}(1) + (1-\eta_{1t}) n_{1t}(0)] + (1-\lambda_t)[\eta_{0t} n_{0t}(1) + (1-\eta_{0t}) n_{0t}(0)]} \tag{7}$$

(g) *Transversality*: dynasties should not be able to access Ponzi-schemes involving continuation values that diverge to plus or minus infinity:

$$\lim_{t\to\infty} \delta^t \Pi_{s=0}^t n_s^\theta V_{j,t} = 0 \tag{8}$$

A *steady state* has the additional feature that all time subscripts can be dropped from the definition above: wages, continuation values, skill ratios and fertility choices must all be stationary (though, to be sure, the aggregate population might change over time). We also require that output be positive. In a steady state, then,

$$\lambda = \frac{\lambda\eta_1 n_1(1) + (1-\lambda)\eta_0 n_0(1)}{\lambda[\eta_1 n_1(1) + (1-\eta_1)n_1(0)] + (1-\lambda)[\eta_0 n_0(1) + (1-\eta_0)n_0(0)]} > 0. \tag{9}$$

The positive output requirement means that we ignore the trivial and uninteresting configuration in which there are no skilled people, there is a huge (infinite) skill premium, and yet the unskilled do not acquire any skills because their wages are zero.[26]

A steady-state proportion of skills can be characterized as follows. For each λ, and given the attendant wages $w_1(\lambda)$ and $w_0(\lambda)$, define $V_0(\lambda)$ and $V_1(\lambda)$ as the unique solutions to the following conditions (which by virtue of U4 generates a contraction mapping from continuation values to current values):

$$V_i(\lambda) = \max_n [u(w_i(\lambda) - r_i(w_i(\lambda))n) + \delta n^\theta V_i(\lambda)]. \tag{10}$$

It is easy to see that $V_1(\lambda)$ decreases in λ while $V_0(\lambda)$ is increasing, that $V_1(\lambda)$ exceeds $V_0(\lambda)$ for low enough values of λ, while the opposite inequality is true at higher values (say for $\lambda \geq \bar{\lambda}$).

Provisionally, think of the $V_i(\lambda)$ defined in this way as the continuation values for children in each skill category. They can't always be the "true" continuation values, as parents may want to switch categories, but in steady state this interpretation will be exactly correct, as non-switching of categories must always be optimal.

Given these values, there exists a parental income threshold $w^*(\lambda)$ at which a parent is just indifferent between imparting skills to all her progeny, or leaving them all unskilled. If parents prefer to not educate their children at every wage rate, set this threshold to zero; in the converse case where parents prefer to educate their children at every wage rate set it equal to ∞. From part (b) of Proposition 1, we know that $w^*(\lambda)$ is uniquely defined. Moreover, if parental wage strictly exceeds $w^*(\lambda)$, the parent has a strict preference for skilled children, while if it is strictly less, she has a strict preference for unskilled children. It follows that a

[26]This configuration is always an equilibrium if $r_0(0) > 0$.

necessary condition for $\lambda > 0$ to be a steady-state skill proportion is

$$w_0(\lambda) \leq w^*(\lambda) \leq w_1(\lambda),$$

with, of course, at least one of these inequalities holding strictly.[27] The unskilled and skilled wage must lie on different sides of the threshold, because it must be optimal for parents to select their own occupation for their children. To understand this, note that (a) competitive equilibrium must be associated with some parents choosing to educate their children, and others that choose not to (to ensure there will be a positive supply into both occupations in the next generation); (b) educational incentives are rising in the wage, and skilled parents must earn a higher wage (in order to induce supply into the skilled occupation).

But this condition is not sufficient when fertility is endogenous: steady state requires checking that the skill ratio is unchanged from one generation to the next. If fertility differs between skilled and unskilled adults, and every adult selects its own occupation for its children, the skill ratio cannot remain constant across generations. If skilled adults have fewer children, for instance, the skill ratio will decline owing to the greater supply into the unskilled occupation in the next generation.[28] Hence if neither wage equals the threshold w^*, and both skilled and unskilled parents strictly prefer their own occupation for their children, an added condition for steady state is that fertility must be the same across skilled and unskilled households. Conversely, if fertility varies with wages, parents in at least one occupation must be indifferent between educating and not educating their children, so as to maintain constancy of the aggregate skill ratio. *Differential fertility must therefore co-exist with mobility.* The pattern of mobility and fertility differentials must be related, as expressed by the following observation.

LEMMA 2: *A skill proportion $\lambda > 0$ is part of a steady state if and only if*

$$w_0(\lambda) \leq w^*(\lambda) \leq w_1(\lambda), \tag{11}$$

with at least one of these inequalities strict, and:

(a) *If $w_0(\lambda) = w^*(\lambda)$, then $n_0(0, \lambda) \geq n_1(1, \lambda)$.*
(b) *If $w^*(\lambda) = w_1(\lambda)$, then $n_0(0, \lambda) \leq n_1(1, \lambda)$.*
(c) *If $w_0(\lambda) < w^*(\lambda) < w_1(\lambda)$, then $n_0(0, \lambda) = n_1(1, \lambda)$.*

In case (a), the unskilled parents are at the threshold w^*, thus indifferent between educating and not educating their children. Since the latter option is associated with a higher fertility, some unskilled parents must decide to educate their

[27] After all, if $w_0(\lambda) = w_1(\lambda)$, then $w^*(\lambda) = \infty$.

[28] More precisely, suppose that both inequalities hold strictly. Then $\eta_1 = 1$ and $\eta_0 = 0$, so (9) implies that $\lambda = \frac{\lambda n_1(1,\lambda)}{\lambda n_1(1,\lambda)+(1-\lambda)n_0(0,\lambda)}$ where $n_i(j, \lambda)$ is the optimally chosen fertility by a parent in category i under the assumption that her children go to category j. This equality calls for the additional requirement that $n_0(0, \lambda) = n_1(1, \lambda)$.

children to ensure the skill ratio in the economy remains constant. So the steady state involves *upward mobility* — the children of some unskilled parents move up to the skilled occupation. Mobility can operate only in one direction: skilled parents must be earning a wage above the threshold and will thus all select their children for the skilled occupation. In this case the cross-sectional fertility-wage correlation must be negative: the upward flow from the unskilled to the skilled occupation must compensate for higher fertility among the unskilled and arrest the resulting downward drift in the skill ratio. The converse is true in case (b): the fertility-wage correlation is positive and there is downward mobility. Hence the model relates the direction of mobility flows to the sign of the fertility-wage correlation. This is verified in the following Proposition, which shows that the average fertility must be declining (rising) in wages if the steady state involves upward (resp. downward) mobility.

PROPOSITION 3: *Only three kinds of steady states are possible:*

(a) *A steady state with upward mobility, in which* $n_1(1) < n_0(0)$*, with* $\eta_1 = 1$ *and* $\eta_0 > 0$*. Average fertility must be declining in wages in any such steady state:*

$$\eta_1 n_1(1) + (1-\eta_1)n_1(0) < \eta_0 n_0(1) + (1-\eta_0)n_0(0). \tag{12}$$

(b) *A steady state with downward mobility, in which* $n_1(1) > n_0(0)$*, with* $\eta_1 < 1$ *and* $\eta_0 = 0$*. Such a steady state must involve average fertility rising in wages, i.e., the opposite of* (12) *must hold.*

(c) *A steady state with no mobility, in which* $n_1(1) = n_0(0)$*, with* $\eta_1 = 1$ *and* $\eta_0 = 0$*. Such a steady state must involve fertility not varying with wages at all.*

That the model may generate occupational mobility in the absence of any stochastic shocks is an interesting by-product of the model. While downward mobility from skilled to unskilled occupations are relatively rare and perhaps explained by bad luck, upward mobility in the reverse direction is usually considered a reflection of determined investments made by the poor and accordingly are held as an indication of equality of opportunity in the society in question. Our theory relates the extent of mobility to fertility patterns, rather than "luck" in the realization of incomes or abilities.

Returning to the question of the fertility-wage pattern, we see that explaining a negative cross-sectional fertility-wage correlation is equivalent to explaining upward mobility. Proposition 3 nevertheless says other kinds of steady states could also conceivably arise, with opposite patterns for mobility and fertility-wage correlations that appear 'unusual' from an empirical standpoint. The key question then is: when can we rule out steady states with zero or negative occupational mobility?

Before proceeding to that question, it helps to settle the question of existence of a non-trivial steady state. In this Lemma 2 turns out to be useful. To ensure that the steady state has positive output and skill ratio, however, we need to impose the assumption of positive fixed costs of child-rearing.[29]

PROPOSITION 4: *There exists a steady state with $\lambda > 0$, provided $r_0(0) > 0$.*

Now turn to the main question of interest: can we say more about the nature of steady states, concerning patterns of mobility and fertility? For this purpose, we specialize to the case of constant-elasticity utility functions. Most existing literature on endogenous fertility restricts attention to this class anyway.

We provide conditions under which the occupational-shift effect invariably dominates the traditional preference-based determinants of fertility choice, so that average fertility declines with income:

PROPOSITION 5: *Assume that utility is isoelastic: $u(c) = \frac{c^{1-\rho}}{1-\rho}$ with ρ different from 1. Then every steady state exhibits upward mobility and a negative fertility-wage correlation if either of the following three conditions hold.*

(a) *In the nonnegative utility case where quality and quantity of children are complements in parental preferences ($\theta \in (0, 1)$), we have $0 < \rho \leq 1 - \theta$.*

(b) *In the negative utility case where quality and quantity of children are substitutes in parental preferences ($\theta < 0$), we have $\rho > 1 - \theta$.*

(c) *The cost structure is separable, and involves time-costs alone; child-rearing costs take the form $k(w) = \psi w$ for some $\psi > 0$.*

While conditions (a)–(c) do not cover all the possible cases under isoelastic utility, Proposition 5 exhibits a broad range of parameter values for which steady states must involve upward mobility. Range (a) covers a region where the result is perhaps to be expected, in which substitution effects associated with parental wage increases are large relative to wealth effects, so the "preference" and "occupational shift" effects go in the same direction.

More surprising are parts (b) and (c), which allow wealth effects to outweigh substitution effects to an arbitrary degree. The preference effect would then be associated with a positive fertility-wage correlation, running counter to the negative occupational shift effect. Nevertheless, under the stated parameter restrictions, the skill premium in wages can be restricted sufficiently to imply that the occupational shift effect must dominate. Such restrictions are imposed by incentive compatibility conditions that steady states must satisfy: if skill premia are too large, excessive incentives would be created for adults to educate their children, which would lead to a shortage of people in the unskilled occupation in the next generation. Such an oversupply into the skilled occupation would cause the skill premium to shrink.

[29]In the absence of such fixed costs, a steady state exists but may involve zero output and skill ratio. Whether an interior steady state can be shown to exist in the absence of this assumption remains an open question.

Nevertheless, we utilise the stated parametric assumptions to obtain a restriction on the skill premium stringent enough to enable the occupational shift effect to dominate the preference effect. Whether the result extends more generally is an open question; we have not yet succeeded in finding a counter-example.

IV. Steady State Determinacy and Comparative Statics

We now turn to the question of local determinacy of steady states. Local determinacy permits us to derive the effects of changed policies, but is of substantive intrinsic interest as well. It bounds the extent of hysteresis or history-dependence that the model permits. In contrast, most models of occupational choice with a discrete set of occupations and exogenous fertility are characterized by a continuum of steady states. We show that incorporation of endogenous fertility into the model eliminates this indeterminacy.

The intuitive basis for this finding is that steady states with either upward or downward mobility can no longer be associated with *strict* incentives for members of *both* occupations: at least one occupation must be indifferent between preparing their children for the same occupation and switching to the alternative occupation. This indifference ties down the steady state skill ratio and per-capita income. And if a steady state involves no mobility, it requires equality of fertility decisions across the two occupations, which also ties down relative wages and hence the equilibrium skill ratio.

To show this formally, introduce a parameter ν of the costs of educating children, and suppose that costs of children prepared for the unskilled occupation $r_0(w)$ is independent of ν, while the costs of children $r_1(w;\nu)$ trained for the skilled occupation is strictly (and smoothly) increasing in ν.

PROPOSITION 6: *Skill ratios forming a steady state with upward or downward mobility are locally unique and finite in number, for a set of parameter values ν of full Lebesgue measure. The same is true for steady states with zero mobility, provided θ is positive.*

The local determinacy of steady states permits us to explore the long run effects of varying costs of child-rearing and of education, as well as regulations pertaining to child labor and redistributive tax-transfer policies. It will be helpful to restrict attention to a linear formulation of child-rearing cost:

$$r_0(w) = f + \psi w, r_1(w) = f + \psi w + x \tag{13}$$

where f denotes the fixed 'goods cost' incurred per child, k the parental time away from work, and x a fixed cost of education.

Moreover, we focus attention on steady states with upward mobility, i.e., on the cases covered by Proposition 5. Note that both functions $w^*(\lambda), w_0(\lambda)$ are increasing in λ. In what follows, refer to these as the w^* and w_0 curves respectively. w^* tends to a negative number as λ tends to 0, and to ∞ as λ tends to $\bar{\lambda}$. At the same time w_0 tends to 0 and $w(\bar{\lambda})$ respectively. Hence there exists at least one skill ratio where w^* and w_0 are equalized, where the w^* curve cuts the w_0 from below (i.e., has a steeper slope). If $n_0 > n_1$, call this a *locally stable* steady state with downward mobility. Intuitively, if λ falls slightly below the steady state, w^* is smaller than w_0. Then both unskilled and skilled households would want to educate their children so the skill ratio will tend to rise. Conversely, if λ rises slightly above the steady state, w^* would be higher than w_0, thus eliminating the willingness of some unskilled households to educate their children, and the skill ratio will fall.

The linear formulation of child rearing costs allows us to obtain a closed form expression for the threshold wage:

$$w^*(\lambda) = \frac{1}{\psi}[x\{(\frac{V_1(\lambda)}{V_0(\lambda)})^{\frac{1}{\theta}} - 1\}^{-1} - f] \tag{14}$$

from the first-order condition (2) for fertility choice.[30] The threshold wage depends on the skill ratio via the dependence of the wage in occupation i and continuation values V_i on this ratio:

$$V_i(\lambda) = \frac{u(w_i(\lambda) - n_i(w_i(\lambda))(f + \psi w_i(\lambda) + xi))}{1 - \delta n_i(w_i(\lambda))^{\theta}}, \tag{15}$$

where in addition $n_i(w_i)$ denotes the optimal fertility choice of a parent with wage w_i and selecting the same occupation for her children.

Small perturbations in child-rearing cost parameters f, x can be shown to be shift the w^* function (and hence the steady state skill ratio) in opposite directions, provided we impose an additional mild assumption on preferences in the case of negative utility):

PROPOSITION 7: *If $\theta < 0$, assume in addition to [U1]–[U4] that $\frac{u'}{-u}$ is decreasing.*[31] *Consider any steady state skill ratio with upward mobility which is locally stable. A small increase (resp. decrease) in fixed cost component f of child-rearing (resp. education cost x) will cause the steady state skill ratio to fall (resp. rise).*

This comparative static result follows from the effect of the parametric changes on parental incentives to invest in the education of their children. A rise in the 'goods cost' of child-rearing increases this incentive, just as a fall in education costs does. The former induces a reduction in fertility, which in turn stimulates

[30]Using E to denote total expenditures $((\psi w + f + xi)n)$ on children, the first order condition implies $E^{1-\theta}U'(w - E) = [\frac{1}{\psi w + f + xi}]^{\theta} V_i$ for occupation choice $i = 0, 1$. This generates condition (14).

[31]This condition is satisfied by constant elasticity utility functions, as well as $u(c) = -exp(-ac)$ with $a > 0$.

an increase in desired quality of children. This is for both a direct reason (child-care expenses per child are lower when not investing in education, so a rise in f raises child-care costs by more for non-investors) and an indirect reason (the continuation utility of skilled children falls by less than for unskilled children).

The model thus 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 children. Conversely, social changes that cause a shift from joint to nuclear families thus induce higher education. Policies of subsidized childcare undermine skill accumulation incentives, and raise inequality between skilled and unskilled wages.

Effects on aggregate fertility are ambiguous. Consider the case of positive θ. A rise in f tends to lower fertility among both skilled and unskilled households at any given skill ratio.[32] This is further reinforced by the induced rise in the proportion of skilled households, since the skilled tend to have fewer children. If wealth effects dominate substitution effects, there is a counteracting effect: fertility within unskilled households rise as a consequence of the rise in the skill ratio.[33] So the net effect on the fertility of the unskilled is ambiguous, while fertility among the skilled must fall. Since the effects on the fertility differential between skilled and unskilled are ambiguous, so are the effects on mobility. It is therefore possible that lower education costs actually end up lowering mobility, if wages and fertility among the unskilled rise sufficiently.[34]

A simple extension of the model allows us to incorporate child labor and study the impact of child labor regulations. 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 pertains to the case where $\gamma = 0$.

Household consumption corresponding to parental wage w is now $c \equiv w - [\psi w + f + xi - \gamma w_0(1 - i)]n$. 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

[32]This is both because of the direct effect of rising child-rearing costs, as well as the induced reduction of continuation values of skilled and unskilled in the case where $\theta > 0$. If $\theta < 0$, the falling continuation values would induce higher fertility. So the ambiguity is even more pronounced in this case.

[33]If the substitution effects dominate instead, fertility among the skilled will rise as a consequence of the fall in their wages. And fertility among the unskilled will fall as their wages rise. In this case the fertility differential will widen, implying a rise in mobility. But the effects on aggregate inequality remain ambiguous.

[34]This provides a potential explanation for the empirical findings of Checchi, Ichino and Rustichini (1999) that mobility in Italy appears to be lower than in the US, despite a more extensive public schooling system and a lower skill premium in wages.

cost. Proposition 7 then implies that both of these induce a rise in the long run 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 a steady state with upward mobility, since it is characterized by intersection of the w^* curve and the w_0 curve. If the steady state is locally stable, this effect raises the steady state skill ratio even further.

Hence the net effect of stronger regulations on child labor is to raise the 'level' of long run development: higher per capita skill and income, with lower wage inequality between the skilled and unskilled. Effects on aggregate fertility and mobility are, however, ambiguous for the reasons explained above.

The model can also be used to study the long run impact of different forms of income redistribution programs. We provide a brief outline of the analysis 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 steady state, 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:

$$\sigma = \frac{\lambda}{1-\lambda}\tau w_1(\lambda) \tag{16}$$

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

$$V_0 = \max_{n_0}[u(w_0 + \sigma - (\psi w_0 + f)n_0) + \delta n_0^\theta V_0] \tag{17}$$

$$V_1 = \max_{n_1}[u((1-\tau)w_1 - (\psi w_1(1-\tau) + f + x)n_1) + \delta n_1^\theta V_1]. \tag{18}$$

These policies lower the value of being skilled, and raise the value of being unskilled. Investment in education is thereby discouraged: the w^* curve shifts up.[35] The long run effect will be to lower per capita skill in the economy. This will raise the skill premium in wages, so the market will undo some of the redistribution.[36]

The adverse long-run effects of unconditional income supports to the poor can

[35] In this case the thresholds differ between skilled and unskilled parents. What matters, of course, in a steady state with upward mobility is the threshold for unskilled parents, and it is this threshold that we refer to here.

[36] Implications for fertility are complex. The wealth effects associated with the transfers will tend to raise fertility among the unskilled and lower them among the skilled. Countering this in the opposite direction are the effects of wage movements induced by the policies, since unskilled wages fall and skilled wages rise.

be avoided with conditional transfers. An example is an education subsidy which is funded by income taxes on the skilled. In this case the continuation value of the unskilled is not directly affected: the value can be computed on the basis of the assumption that they do not invest in education of their children, whence they receive no benefits from the transfers. On the other hand, the continuation value of being skilled rises, since skilled households have more options.[37] This encourages investment in education: the w^* curve shifts down, and the steady state skill ratio rises (hence so does per capita income, while the skill premium declines). The effects are exactly the opposite of an unconditional welfare system.

V. Related Literature

We start by describing related literature on the wage-fertility correlation.

First, as discussed in the introduction, there is the view that the cross-sectional relationship is indeed positive, and the negative relationship we do see in most empirical studies is the result of some omitted variable. Becker (1960) is a proponent of this point of view, emphasizing the possible differences between *desired* and *actual* numbers of children, owing to ignorance concerning contraceptive methods. Another important omitted variable (see, e.g., Freedman 1963) is that family income is correlated with greater female participation in the labor force, and it is the latter that drives the decline. (This view can, of course, be folded into the substitution effect which is driven by time costs of rearing children.) Under this view, then, theory has little to say about the net effect.

Second, there is the view that the cross-sectional relationship is "truly" negative, and must be explained by the theory at hand. Attempts include appropriate calibration of the parameters, so that substitution effects kill off the income effect, introduction of 'quality' of children as an additional choice coupled with suitable assumptions concerning child-rearing costs (Becker and Lewis 1973; Becker and Tomes 1976; Moav 2005).[38] An alternative approach is to introduce non-

[37] An education subsidy π lowers private education cost to $x - \pi$. Hence budget balance requires $\tau w_1(\lambda) = n_1(\lambda)\pi$, i.e., skilled households pay the taxes to fund the education subsidies they would avail if they chose the same fertility as in the absence of the subsidy. In that case they are as well off as before. But they have the option to have a different number of children, which could make them better off.

[38] As Jones-Schoonbroodt-Tertilt (2008, Section 5.2) explain, this model assumes goods costs of children are $b_0 + b_1 w$, and schooling expenditures s determines child quality or education linearly through $e = d_0 + d_1 s$. If we specialize to the case of two levels of education $e = 0, 1$ this reduces to $r_0(w) = b_0 + b_1 w, r_1(w) = b_0 + b_1 w + \frac{e}{d_1} - \frac{d_0}{d_1})$. This forms a special case of our model, as these functions satisfy R1–R3. In partial equilibrium with a Cobb-Douglas preference function incorporating parental consumption, fertility and child education as separate arguments (in an additively separable manner), which represents the Becker-Tomes-Moav model, a negative wage-fertility correlation obtains only if $b_0 < \frac{d_0}{d_1}$, i.e., the net goods cost per child is negative. This requires the 'innate ability' of children to be large enough to completely offset the goods cost of childcare, a very restrictive assumption. Our partial equilibrium analysis differs not only in terms of the less restrictive assumptions on childcare costs, but also in incorporating the Barro-Becker formulation of parental altruism which creates a nonseparability between quality and quantity of children in parental preferences. This in turn creates nonconvexities and non-interior solutions for quality choices, as we explain in Section 3. So even if we were to incorporate the restrictive Becker-Tomes-Moav assumptions on childcare costs, their conclusions would not apply.

homotheticities in preferences (see, e.g. Galor and Weil 2000; Greenwood and Seshadri 2002; or Fernández, Guner and Knowles 2005).

Jones, Schoonbroodt and Tertilt (2008) provide an overview of these explanations, and argue that they all require restrictive assumptions. Instead they suggest an argument that relies on heterogeneity in parental tastes. Parents who want larger families will realize that more time will be spent on bringing up children, and this lowers their optimal choice of human capital, not for their children as in our framework, but for *themselves*. Therefore larger families could be correlated with lower family income.

Steady state models with endogenous fertility, in which the cross-sectional relationship is explicitly addressed, include Alvarez (1999) and Kremer and Chen (1999, 2002). Alvarez introduces Barro-Becker preferences into Loury's (1981) theory of intergenerational inequality. The main objective is to see how the persistence result in inequality (arising from shocks) withstands endogenous fertility, but as a byproduct Alvarez obtains the result that fertility is *positively* correlated with parental wealth.[39] In a modification of the model, Alvarez introduces various non-separabilities in preferences and rearing costs; in particular, increasing marginal costs in child-rearing. If these effects are strong enough, this baseline finding may be reversed. In summary, the model delivers the opposite result on fertility-wealth in its baseline setup and an ambiguous prediction when the modifications are introduced.

Kremer and Chen use a special specification of parental preferences, in which utility is given by $c + \log n$. Parents must spend time in child-rearing, but do not care about the quality of children or the payoffs enjoyed by them. This assumed absence of wealth effects implies that rising parental wages exert only substitution effects on the demand for fertility. Hence fertility is decreasing in parental wages. In this sense, the Kremer-Chen exercise is a special case of the parametric calibration that forces substitution effects to dominate income effects.

The same is true of Doepke (2004), who assumes that utility functions exhibit lower curvature than the log function, and that there are no goods costs involved in child-rearing. In this case, as we have seen in Proposition 2, wealth effects are weaker than substitution effects, ensuring a negative wage-fertility correlation. However some of our partial equilibrium findings are related to results in Doepke (2004). For instance his Proposition 3 implies the jump-down as occupational boundaries are crossed (derived for a more restrictive class of preferences and rearing/schooling costs). Doepke's Proposition 4 makes the point that in equilibrium, there could be upward, downward or zero mobility outcomes, and also observes that in the first two cases the parents in the "out-migrating" category will need to be indifferent between educating and not educating their children. Doepke subsequently expresses the view that the first of these three outcomes is the most relevant one, and then moves on to calibration and empirical investi-

[39]In fact, in his baseline model, the positive correlation is so strong that per-child parental transfers are independent of parental wealth.

gation. Our contribution is to provide a theoretical justification for the focus on this class of steady states. We additionally provide theoretical results concerning local determinacy and comparative static properties of steady states.

Finally, we discuss the connections to the occupational choice literature. The generic model of occupational choice with fixed fertility takes as given a set of occupations (typically two, as in the exercise here).[40] The steady-state conditions are simple: no dynasty currently in one occupation must wish to move their progeny to another with a different setup cost. This occupational persistence follows from single-crossing, which rules out cycles in occupational choice. In particular, there is no mobility in steady state. This also implies a continuum of steady states with a finite number of occupations.[41] Such indeterminacy makes it difficult to analyze the long-run impact of policy changes.

VI. Concluding Remarks

We have developed a theory on the implications of interactions between fertility and human capital in a setting with imperfect financial markets. This framework provides new insights into the wage-fertility relationship, the determinants of intergenerational mobility, and the extent of macroeconomic history dependence. It provides a tractable comparative static analysis of long-run effects of changes in a variety of fiscal and human capital policies. It also illustrates a number of possible factors underlying observed features of the demographic transition in developing countries, whereby economic and social factors associated with urbanization and modernization induce large declines in fertility (e.g., falling costs of education, rising costs of childcare, regulations on child labor).

Our approach is based on a general equilibrium argument, using the discipline imposed by a steady state which limits the extent to which skilled wages exceed unskilled wages, and thus restrict the scope of a positive (net) correlation arising from preferences alone. Our results are entirely analytical and do not rely on numerical computations; they are robust to a substantially wider range of assumptions concerning parental preferences than most of the existing literature. Proposition 5 is clearly very different from all of the preceding explanations of the wage-fertility correlation, as it does not rest on assumed weakness of wealth effects or unobserved preference heterogeneity, but instead on the endogeneity of occupational wages and related steady state restrictions.

Future research is needed to address the question of how our model could be extended to more than two occupations. We have worked out extensions of some of the main results to a finite number of occupations. Even within the context of

[40]Versions of this model appear in Banerjee and Newman (1993), Ljunqvist (1993), Galor and Zeira (1993), Ghatak and Jiang (2002), and Mookherjee and Ray (2003). Ray (1990, 2006) contains a model which exactly aligns with this generic description in the two-occupation case.

[41]Mookherjee and Ray (2003) show that a continuum of occupations restores uniqueness. Mookherjee and Napel (2007) show that if there are stochastic shocks to "ability," then steady states are locally isolated.

the two occupation model some gaps in our analysis remain: for instance whether Proposition 5 extends to a wider set of preferences and technology.

We hope our results will spur efforts to test the central predictions of our model. The central finding of this paper is that the occupational shift effect accounts for a robust negative correlation between parental wages and fertility, which may work against a positive wealth effect. Hence *within* occupations or human capital categories, or in contexts where there is not much scope for occupational or human capital variations, the wage-fertility correlation may be positive. Similarly over short periods of time when occupations of the adult population are unchanging, fertility may move in the same direction as income. But the correlation *across* occupational categories will typically be negative. These are consistent with empirical findings reported in some earlier literature such as Freedman (1963) and Simon (1969), and it would be interesting to see if they continue to be confirmed more broadly using more powerful econometric techniques and other datasets (e.g. pertaining to developing countries).

REFERENCES

Aghion, Philippe, and Patrick Bolton. 1997. “A Theory of Trickle-Down Growth and Development.” *Review of Economic Studies,* 64(2), 151–172.

Alvarez, Fernando. 1999. “Social Mobility: The Barro-Becker Children Meet the Laitner-Loury Dynasties.” *Review of Economic Dynamics,* 2(1), 65–103.

Banerjee, Abhijit V., and Andrew F. Newman. 1993. “Occupational Choice and the Process of Development.” *Journal of Political Economy,* 101(2): 274–298.

Barro, Robert J., and Gary S. Becker. 1989. “Fertility Choice in a Model of Economic Growth.” *Econometrica,* 57(2): 481–501.

Basu, Kaushik, and Pham Hoang Van. 1998. “The Economics of Child Labor.” *American Economic Review,* 88(3): 412–27.

Becker, Gary S. 1960. “An Economic Analysis of Fertility.” In *Demographic and Economic Change in Developed Countries*, ed. G. S. Becker, 209–240. Princeton: Princeton University Press.

Becker, Gary S., and Robert J. Barro. 1986. “Altruism and the Economic Theory of Fertility.” *Population and Development Review,* 12, Supplement Issue, 69–76.

Becker, Gary S., and Robert J. Barro. 1988. “A Reformulation of the Economic Theory of Fertility.” *Quarterly Journal of Economics,* 103(1): 1–25.

Becker, Gary S., and H. Gregg Lewis. 1973. “On the Interaction between the Quantity and Quality of Children.” *Journal of Political Economy,* 81(2): S279–S288.

Becker, Gary S., and Nigel Tomes. 1976. “Child Endowments and the Quantity and Quality of Children.” *Journal of Political Economy,* 84(4): S143–S162.

Becker, Gary S., and Nigel Tomes. 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy,* 87(6): 1153–89.

Blau, David, and Wilbert van der Klaauw. 2007. "The Impact of Social and Economic Policy on the Family Structure Experiences of Children in the United States." Unpublished.

Checchi, Daniele, Andrea Ichino, and Aldo Rustichini. 1999. "More Equal but less Mobile? Education Financing and Intergenerational Mobility in Italy and in the U.S." *Journal of Public Economics,* 74(3): 351–393.

Clark, Gregory. 2005. "Human Capital, Fertility, and the Industrial Revolution." *Journal of the European Economic Association, Papers and Proceedings* 3(2-3): 505–515.

Clark, Gregory. 2007. *Farewell to Alms: A Brief Economic History of the World,* Princeton University Press: Princeton.

Clark, Gregory, and Gillian Hamilton. 2006. "Survival of the Richest: The Malthusian Mechanism in Pre-Industrial England," *Journal of Economic History,* 66(3): 707–736.

Dahan, Momi, and Daniel Tsiddon. 1998. "Demographic Transition, Income Distribution and Economic Growth." *Journal of Economic Growth,* 3(1): 29–52.

De La Croix, David, and Matthias Doepke. 2003. "Inequality and Growth: Why Differential Fertility Matters." *American Economic Review,* 93(4): 1091–1113.

Doepke, Matthias. 2004. "Accounting for Fertility Decline During the Transition to Growth." *Journal of Economic Growth,* 9(3): 347–383.

Doepke, Matthias. 2005. "Child Mortality and Fertility Decline: Does the Barro-Becker Model Fit the Facts?" *Journal of Population Economics,* 18(2): 337–366.

Doepke, Matthias, and Fabrizio Zilibotti. 2005. "The Macroeconomics of Child Labor Regulation." *American Economic Review,* 95(5): 1492–1524.

Easterlin, Richard. 1973. "Relative Economic Status and the American Fertility Swings." In *Family Economic Behavior*, ed. E. B. Sheldon, 170–223. Lippincott: Philadelphia.

Fernández, Raquel, Nezih Guner, and John Knowles. 2005. "Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality." *Quarterly Journal of Economics,* 120(1): 273–344.

Freedman, Deborah S. 1963. "The Relation of Economic Status to Fertility." *American Economic Review,* 53(3): 414–426.

Freeman, Scott. 1996. "Equilibrium Income Inequality among Identical Agents." *Journal of Political Economy,* 104(5): 1047–1064.

Galor, Oded, and Zeira, Joseph. 1993. "Income Distribution and Macroeconomics." *Review of Economic Studies,* 60(1): 35–52.

Galor, Oded, and David N. Weil. 2000. "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond." *American Economic Review,* 90(4): 806–828.

Ghatak, Maitreesh, and Neville N.-H. Jiang. 2002. "A Simple Model of Inequality, Occupational Choice and Development." *Journal of Development Economics,* 69(1): 205–226.

Greenwood, Jeremy, and Ananth Seshadri. 2002. "The U.S. Demographic Transition." *American Economic Review, Papers and Proceedings,* 92(2): 153–159.

Haines, Michael R. (1976. " Population and Economic Change in Nineteenth Century Eastern Europe: Prussian Upper Silesia, 1840-1913." *The Journal of Economic History*, 36(2): 334–358.

Jones, Larry E., and Alice Schoonbroodt. 2010. "Complements versus Substitutes and Trends in Fertility Choice in Dynamic Models" *International Economic Review,* 51(3): 671–699.

Jones, Larry E., Alice Schoonbroodt, and Michele Tertilt. 2011. "Fertility Theories: Can They Explain the Negative Fertility-Income Relationship?" In *Demography and the Economy,* ed. J. B. Shoven, 43–100. University of Chicago Press: Chicago.

Jones, Larry E., and Michele Tertilt. 2008. "An Economic History of Fertility in the United States: 1826–1960." In *Frontiers of Family Economics*, ed. P. Rupert, Vol. 1, 165–230, Emerald Press.

Kremer, Michael, and Daniel L. Chen. 1999. "Income Distribution Dynamics with Endogenous Fertility." *American Economic Review,* 89(2): 155–160.

Kremer, Michael, and Daniel L. Chen. 2002. "Income Distribution Dynamics with Endogenous Fertility." *Journal of Economic Growth,* 7(3): 227–258.

Lee, Donald R. 1987. "Population Dynamics of Humans and Other Animals." *Demography,* 24(4): 443–466.

Ljungqvist, Lars. 1993. "Economic Underdevelopment: The Case of Missing Market for Human Capital." *Journal of Development Economics,* 40(2): 219–239.

Lloyd-Ellis, Huw, and Dan Bernhardt. 2000. "Enterprise, Inequality and Economic Development," *Review of Economic Studies,* 67(1): 147–168.

Loury, Glenn C. 1981. "Intergenerational Transfers and the Distribution of Earnings." *Econometrica,* 49(4): 843–867.

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green. 1995. *Microeconomic Theory,* New York: Oxford University Press.

Matsuyama, Kiminori. 2000. "Endogenous Inequality." *Review of Economic Studies,* 67(4): 743–759.

Matsuyama, Kiminori. 2006. "The 2005 Lawrence Klein Lecture: Emergent Class Structure," *International Economic Review,* 47(2): 327–260.

Moav, Omer. 2005. "Cheap Children and the Persistence of Poverty." *Economic Journal,* 115(500): 88–110.

Mookherjee, Dilip, and Stefan Napel. 2007. "Intergenerational Mobility and Macroeconomic History Dependence," *Journal of Economic Theory,* 137(1): 49–78.

Mookherjee, Dilip, and Debraj Ray. 2002. "Is Equality Stable?" *American Economic Review, Papers and Proceedings,* 92(2): 253–259.

Mookherjee, Dilip, and Debraj Ray. 2003. "Persistent Inequality," *Review of Economic Studies,* 70(2): 369–394.

Mookherjee, Dilip, and Debraj Ray. 2008. "A Dynamic Incentive-Based Argument for Conditional Transfers." *Economic Record,* 84(S1): S2–16.

Ray, Debraj. 1990. "Income Distribution and Macroeconomic Behavior." Unpublished.

Ray, Debraj. 2006. "On the Dynamics of Inequality." *Economic Theory,* 29(2): 291–306.

Simon, Julian L. 1969. "The Effect of Income on Fertility." *Population Studies*, 23(3): 327–341.

Simon, Julian L. 1977. *The Economics of Population Growth,* Princeton University Press: Princeton, New Jersey.

Spengler, Joseph J. 1952. "Population Theory." In *A survey of Contemporary Economics, Vol. 2* ed. B. Haley, 99–106. American Economic Association: Homewood.

Wrigley, Edward A. 1961. *Industrial Growth and Population Change,* Cambridge University Press: Cambridge.

Appendix: Proofs

PROOF OF PROPOSITION 1. (a) For any given value of w, set $x \equiv r_1(w) - r_0(w) > 0$. Differentiate the first line of (3) with respect to e and use the envelope theorem to get

$$\frac{\partial V(w,e)}{\partial e} = n(w,e)^{\theta}(V_1 - V_0) - u'(w - r(w,e)n(w,e))xn(w,e),$$

and now use the first-order condition (2) to write this as

$$\text{(A1)} \qquad \frac{\partial V(w,e)}{\partial e} = \frac{n(w,e)^{\theta}x}{r(w,e)}\left[\left\{\frac{r_0(w)}{x} + (1-\theta)e\right\}(V_1 - V_0) - \theta V_0\right].$$

Of course, if $V_1 \le V_0$, $e = 0$ is optimal. If $V_1 > V_0$ then (A1) proves that V is strictly quasiconvex in e, for given w. It follows that no interior solution to e can ever maximize V, establishing (a).

Next we prove (c). Combine part (a) with Lemma 1 to conclude the following. Consider any wage w^* at which the parent is indifferent between $e = 0$ and $e = 1$. Then the expenditures on children must be equalized at such a wage:

$$\text{(A2)} \qquad r_1(w^*)n(w^*,1) = r_0(w^*)n(w^*,0).$$

In turn, this implies that $n(w^*,1) < n(w^*,0)$ at a point of indifference.

Now consider (b). Suppose that both $e = 0$ and $e = 1$ are optimal for some parental wage w^*. Define

$$\Delta(w) = u\left(w - r_1(w)n(w,1)\right) + n(w,1)^{\theta}V_1 - u\left(w - r_0(w)n(w,0)\right) - n(w,0)^{\theta}V_0$$

and differentiate with respect to w, evaluating the result at w^* where $\Delta(w^*) = 0$. By the envelope theorem applied to $n(w*,1)$ and $n(w*,0)$, we have that

$$\Delta'(w^*) = u'\left(w^* - r_1(w^*)n(w^*,1)\right)[1 - r_1'(w^*)n(w^*,1)] - u'\left(w^* - r_0(w^*)n(w^*,0)\right)[1 - r_0'(w^*)n(w^*,0)].$$

By the reasoning above, the terms within the two u's are exactly equal. It follows that the sign of $\Delta'(w^*)$ equals the sign of

$$r_0'(w^*)n(w^*,0) - r_1'(w^*)n(w^*,1) = \frac{r_0'(w^*)}{r_0(w^*)}r_0(w)n(w^*,0) - \frac{r_1'(w^*)}{r_1(w^*)}r_1(w^*)n(w^*,1).$$

Using (A2) and applying (R.3), we see that $\Delta'(w^*) > 0$.

PROOF OF PROPOSITION 2. Study the first order condition (2), and note that the right-hand side of this condition is independent of w, while it is strictly decreasing in n. Recall that continuation value V and θ have the same sign, and that $\theta < 1$. Therefore n_i is locally increasing in w if the derivative of the left-hand side of (2) with respect to w is negative, or equivalently, if

$$u'(c)r'(w) + u''(c)r(w)[1 - nr'(w)] < 0,$$

and is decreasing if the opposite inequality holds. After noting that $1 - \psi_i n_i \geq 1 - (f_i + \psi_i)n_i = c_i/w > 0$, it is easy to see that this expression reduces to condition (5).

PROOF OF LEMMA 2. The discussion preceding the Lemma already establishes the necessity of (11), as well as part (c). Parts (a) and (b) are established in similar fashion. For instance, to establish (a), suppose that $w_0(\lambda) = w^*(\lambda)$; then $w^*(\lambda) < w_1(\lambda)$. It follows that skilled parents strictly prefer skilled children, so that $\eta_1 = 1$. Therefore — because $n_i(j) = n_i(j, \lambda)$ for every i and j — (9) implies that

$$\begin{aligned}
\lambda &= \frac{\lambda n_1(1,\lambda) + (1-\lambda)\eta_0 n_0(1,\lambda)}{\lambda n_1(1,\lambda) + (1-\lambda)[\eta_0 n_0(1,\lambda) + (1-\eta_0)n_0(0,\lambda)]} \\
&\geq \frac{\lambda n_1(1,\lambda)}{\lambda n_1(1,\lambda) + (1-\lambda)(1-\eta_0)n_0(0,\lambda)} \\
&\geq \frac{\lambda n_1(1,\lambda)}{\lambda n_1(1,\lambda) + (1-\lambda)n_0(0,\lambda)},
\end{aligned}$$

which implies right away that $n_0(0, \lambda) \geq n_1(1, \lambda)$. Part (b) is established in a parallel way.

To establish sufficiency, pick $\lambda > 0$ such that (11) and one of (a)–(c) are satisfied. Let the associated wages be $w_1 = w_1(\lambda)$ and $w_0 = w_0(\lambda)$ and associated continuation values be $V_1 = V_1(\lambda)$ and $V_0 = V_0(\lambda)$, as given by (10). Let $n_i(j) = n_i(j, \lambda)$ for every i and j. If case (a) applies, we have $n_0(0, \lambda) \geq n_1(1, \lambda)$, so that

$$\lambda \geq \frac{\lambda n_1(1,\lambda)}{\lambda n_1(1,\lambda) + (1-\lambda)n_0(0,\lambda)}.$$

Of course, $w^*(\bar{\lambda}) = \infty$, which means that $\lambda < \bar{\lambda} < 1$. It is therefore easy to see that there exists $\eta_0 \in [0, 1)$ such that

$$\lambda = \frac{\lambda n_1(1,\lambda) + (1-\lambda)\eta_0 n_0(1,\lambda)}{\lambda n_1(1,\lambda) + (1-\lambda)[\eta_0 n_0(1,\lambda) + (1-\eta_0)n_0(0,\lambda)]}.$$

Choose this value of η_0 and set $\eta_1 = 1$, and now check that all conditions for a steady state are satisfied. In particular, (11) guarantees that it is optimal never to switch categories, so that the V_i's represent the true continuation values.

Similar arguments apply for cases (b) or (c).

PROOF OF PROPOSITION 3. A steady state must have either (i) $n_1(1) < n_0(0)$, (ii) $n_1(1) > n_0(0)$, or (iii) $n_1(1) = n_0(0)$. We show that these three cases must respectively correspond to (a)–(c) in the statement of the proposition.

Consider case (a), in which $n_1(1) < n_0(0)$. Then part (a) of Lemma 2 is applicable, so that $w_0 = w^*(\lambda) < w_1$. It follows that $\eta_1 = 1$. We claim, moreover, that $\eta_0 > 0$. Suppose not, then $\eta = 0$, and so, using (9) with $\eta_1 = 1$,

$$\lambda = \frac{\lambda n_1(1)}{\lambda n_1(1) + (1-\lambda)n_0(0)} < \lambda,$$

a contradiction.

Next, we show that average fertility is declining in wages:

$$n_1(1) < \eta_0 n_0(1) + (1-\eta_0)n_0(0). \tag{A3}$$

Using (9) with $\eta_1 = 1$, we have that

$$\begin{aligned}\lambda &= \frac{\lambda n_1(1) + (1-\lambda)\eta_0 n_0(1)}{\lambda n_1(1) + (1-\lambda)[\eta_0 n_0(1) + (1-\eta_0)n_0(0)]} \\ &> \frac{\lambda n_1(1)}{\lambda n_1(1) + (1-\lambda)[\eta_0 n_0(1) + (1-\eta_0)n_0(0)]},\end{aligned}$$

where the inequality uses the facts that $\lambda \in (0,1)$, $\eta_0 > 0$, and $n_i(j) > 0$ for all i and j. Cross-multiplying and transposing terms, we see that

$$\lambda(1-\lambda)[\eta_0 n_0(1) + (1-\eta_0)n_0(0)] > \lambda(1-\lambda)n_1(1)$$

which establishes (A3).

The argument for case (b) is analogous, while (c) is obvious.

PROOF OF PROPOSITION 4. We display $\lambda > 0$ such that (11) and one of the conditions in (a)–(c) of Lemma 2 is met. Observe that $V_1(\lambda)$ is decreasing and continuous in λ, while $V_0(\lambda)$ is increasing and continuous in λ. It is easy to conclude that $w^*(\lambda)$ is continuous in λ (in the extended reals) and that it is strictly increasing as long as it is finite.[42]

On the other hand, $w_1(\lambda)$ is continuous and decreasing in λ, with the assumed end-point conditions.

We must conclude that there exists (unique) $\lambda_1 > 0$ such that $w_1(\lambda_1) = w^*(\lambda_1)$. If at this value, $n_1(1,\lambda_1) \geq n_0(0,\lambda_1)$, we are done (use part (b) of Lemma 2).

Otherwise $n_1(1,\lambda_1) < n_0(0,\lambda_1)$. It is obvious that $n_1(1,\lambda)$ is bounded away from 0 as $\lambda \to 0$ (both parental income and $V_1(\lambda)$ go to infinity). On the other

[42]That is, $w^*(\lambda)$ is increasing and continuous whenever it is finite, $w^(\lambda') = \infty$ if $\lambda' > \lambda$ and $w^*(\lambda) = \infty$, and $w^*(\lambda_n) \to \infty$ if $\lambda_n \to \lambda$ and $w^*(\lambda) = \infty$.

hand, given that $r_0(0) > 0$, it must be that $n_0(0, \lambda) \to 0$. It is easy to see that, moreover, that $n_i(i, \lambda)$ is continuous for $i = 0, 1$. It follows that there exists a *largest* value of λ smaller than λ_1 — call it λ_2 — such that $n_1(1, \lambda_2) \geq n_0(0, \lambda_2)$.

Indeed, by continuity of $n_i(i, \lambda)$, we must have $n_1(1, \lambda_2) = n_0(0, \lambda_2)$. Also, $w^*(\lambda_2) \leq w_1(\lambda_2)$.[43] If $w_0(\lambda_2) \leq w^*(\lambda_2)$ as well, then we are again done (use part (c) of Lemma 2).

Otherwise $w_0(\lambda_2) > w^*(\lambda_2)$. Define λ_3 to be the smallest value of $\lambda > \lambda_2$ such that $w_0(\lambda_3) = w^*(\lambda_3)$. It is obvious that $\lambda_3 \in (\lambda_2, \lambda_1)$.[44] We claim that $n_1(1, \lambda_3) < n_0(0, \lambda_3)$. This follows right away from our definition of λ_2 as the *largest* value of λ smaller than λ_1, such that $n_1(1, \lambda_2) \geq n_0(0, \lambda_2)$. Now the condition in part (a) of Lemma 2 is met, and the proof is complete.

PROOF OF PROPOSITION 5. For ease of notation, denote $n_0(0)$ by n_0 and $n_1(1)$ by n_1. We first establish the following

LEMMA 3: *Suppose that at least one of the conditions in Proposition 5 holds. Then in any steady state,*

$$\frac{w_1}{r_1(w_1)} < \frac{w_0}{r_0(w_0)}. \tag{A4}$$

PROOF:

Suppose that (A4) is false in some steady state. Then $w_1 - r_1(w_1)n_0 \geq 0$,[45] and so, because n_1 is an optimal choice at parental wage w_1,

$$V_1 = \frac{[w_1 - r_1(w_1)n_1]^{1-\rho}}{1-\rho} + \delta n_1^\theta V_1 \geq \frac{[w_1 - r_1(w_1)n_0]^{1-\rho}}{1-\rho} + \delta n_0^\theta V_1.$$

It follows that

$$V_1 \geq \frac{[w_1 - r_1(w_1)n_0]^{1-\rho}}{(1-\rho)(1-\delta n_0^\theta)} \geq \left[\frac{r_1(w_1)}{r_0(w_0)}\right]^{1-\rho} \frac{[w_0 - r_0(w_0)n_0]^{1-\rho}}{(1-\rho)(1-\delta n_0^\theta)} = \left[\frac{r_1(w_1)}{r_0(w_0)}\right]^{1-\rho} V_0. \tag{A5}$$

Now consider a parent with wage w_0. Suppose that she chooses a fertility of $\tilde{n}$, where

$$\tilde{n} r_1(w_0) = n_0 r_0(w_0), \tag{A6}$$

[43] After all, $w_1(\lambda_1) = w^*(\lambda_1)$, the former function is declining in λ, the latter increasing in λ, and $\lambda_2 < \lambda_1$.

[44] After all, at λ_1 we have $w^*(\lambda_1) = w_1(\lambda_1) > w_0(\lambda_1)$, the strict inequality following from the fact that $w^*(\lambda_1)$ is finite.

[45] Because $w_0 - r_0(w_0)n_0 \geq 0$, we have $\frac{w_0}{r_0(w_0)} \geq n_0$. Now use the negation of (A4).

and educates all her children. Then her overall payoff is given by

$$\begin{aligned}
\text{(A7)} \qquad \hat{V}_0 &\equiv \frac{[w_0 - r_1(w_0)\tilde{n}]^{1-\rho}}{1-\rho} + \delta\tilde{n}^{\theta} V_1 \\
\text{(A8)} \qquad &\geq \frac{[w_0 - r_1(w_0)\tilde{n}]^{1-\rho}}{1-\rho} + \delta\tilde{n}^{\theta} \left[\frac{r_1(w_1)}{r_0(w_0)}\right]^{1-\rho} V_0 \\
\text{(A9)} \qquad &= \frac{[w_0 - r_0(w_0)n_0]^{1-\rho}}{1-\rho} + \delta {n_0}^{\theta} \left[\frac{r_0(w_0)}{r_1(w_0)}\right]^{\theta} \left[\frac{r_1(w_1)}{r_0(w_0)}\right]^{1-\rho} V_0,
\end{aligned}$$

where the first inequality follows from (A5), and the last equality from (A6).

Now, if there is positive utility and $1-\rho \geq \theta$, as assumed, then
(A10)

$$\left[\frac{r_0(w_0)}{r_1(w_0)}\right]^{\theta} \left[\frac{r_1(w_1)}{r_0(w_0)}\right]^{1-\rho} \geq \left[\frac{r_0(w_0)}{r_1(w_0)}\right]^{1-\rho} \left[\frac{r_1(w_1)}{r_0(w_0)}\right]^{1-\rho} = \left[\frac{r_1(w_1)}{r_1(w_0)}\right]^{1-\rho} > 1,$$

where the first inequality uses R.2, and the last inequality uses R.1.

On the other hand, if there is negative utility and $1-\rho \leq \theta$, as assumed, then
(A11)

$$\left[\frac{r_0(w_0)}{r_1(w_0)}\right]^{\theta} \left[\frac{r_1(w_1)}{r_0(w_0)}\right]^{1-\rho} \leq \left[\frac{r_0(w_0)}{r_1(w_0)}\right]^{1-\rho} \left[\frac{r_1(w_1)}{r_0(w_0)}\right]^{1-\rho} = \left[\frac{r_1(w_1)}{r_1(w_0)}\right]^{1-\rho} \leq 1,$$

where R.1 and R.2 are used again at exactly the same points.

Noting that $V_0 > 0$ in the positive utility case and $V_0 < 0$ in the negative utility case, we can use (A10) or (A11) in equation (A9) (depending on the case we are in) to conclude that

$$\hat{V}_0 > \frac{[w_0 - r_0(w_0)n_0]^{1-\rho}}{1-\rho} + \delta {n_0}^{\theta} V_0 = V_0,$$

which violates part (b) of Proposition 4 for a steady state.

Finally, if condition (c) of Proposition 5 holds, (A4) is obtained free of charge. For

$$\frac{w_1}{r_1(w_1)} = \frac{w_1}{\psi w_1 + x(w_1)} < \frac{1}{\psi} = \frac{w_0}{r_0(w_0)}.$$

Now we turn to the main proof. The two first-order conditions for the choice of n_0 and n_1 tell us that

$$u'(c_0)r_0(w_0)n_0^{1-\theta} = \frac{\theta\delta u(c_0)}{1-\delta n_0^{\theta}} \text{ and } u'(c_1)r_1(w_1)n_1^{1-\theta} = \frac{\theta\delta u(c_1)}{1-\delta n_1^{\theta}}.$$

Using the constant elasticity specification, these equalities imply that

$$\text{(A12)} \qquad \frac{1-\rho}{\theta}\frac{r_0(w_0)}{c_0} = \frac{\delta n_0^{\theta-1}}{1-\delta n_0^{\theta}} \text{ and } \frac{1-\rho}{\theta}\frac{r_1(w_1)}{c_1} = \frac{\delta n_1^{\theta-1}}{1-\delta n_1^{\theta}},$$

and combining,

$$\text{(A13)} \qquad \frac{c_1}{c_0} = \frac{r_1(w_1) n_0^{\theta-1}(1-\delta n_1^{\theta})}{r_0(w_0) n_1^{\theta-1}(1-\delta n_0^{\theta})}.$$

If there is a steady state with zero mobility, so that $n_0 = n_1$, then (A13) immediately implies that

$$\frac{w_1 - r_1(w_1)n_1}{r_1(w_1)n_1} = \frac{c_1}{r_1(w_1)n_1} = \frac{c_0}{r_0(w_0)n_0} = \frac{w_0 - r_0(w_0)n_0}{r_0(w_0)n_0},$$

and using $n_0 = n_1$ once again, we must conclude that

$$\frac{w_1}{r_1(w_1)} = \frac{w_0}{r_0(w_0)},$$

which contradicts the assertion (A4) of Lemma 3.

We now eliminate steady states with downward mobility. The following lemma completes part of this task:

LEMMA 4: *If $\rho+\theta \geq 1$ in the nonnegative utility case, and without any further assumptions in the negative utility case, n_i and $w_i/r_i(w_i)$ must co-move over the two occupations $i = 0, 1$.*

PROOF:

Recall (A12). Define $\alpha \equiv (1-\rho)/\theta$ (always a positive number) and $\mu_i \equiv w_i/r_i(w_i)$ for $i = 0, 1$. Then (A12) can be written as

$$\text{(A14)} \qquad \delta(\alpha - 1)n_i^{\theta} + \delta \mu_i n_i^{\theta-1} = \alpha$$

for $i = 0, 1$. The left hand side of this expression is strictly increasing in μ_i. By using a standard argument, we establish the co-movement of n_i and μ_i if we can show that the derivative of the left hand side in n_i is strictly negative, evaluated at the equality in (A14). To this end, drop the i-subscripts, define

$$D(n) \equiv \delta(\alpha - 1)n^{\theta} + \delta \mu n^{\theta-1}$$

and differentiate with respect to n to see that

$$D'(n) = \delta(\alpha - 1)\theta n^{\theta-1} + \delta \mu(\theta - 1)n^{\theta-2}$$

So we are already done in the nonnegative utility case under the assumption that $\rho + \theta \geq 1$, for then $\alpha \leq 1$ and $\theta < 1$. Otherwise, we are in the negative utility case, and

$$\begin{aligned} nD'(n) &= \theta[\delta(\alpha - 1)n^\theta + \delta\mu n^{\theta-1}] - \delta\mu n^{\theta-1} \\ &= \theta\alpha - \delta\mu n^{\theta-1}, \end{aligned}$$

where the last equality uses (A14). This expression is negative, because $\theta < 0$.

Combining Lemmas 3 and 4, the proof of the proposition is complete under all conditions except (a). In the remainder of the proof, then, we concentrate on the nonnegative utility case with $\rho + \theta \leq 1$.

Suppose, on the contrary, that a steady state displays downward mobility in case (a). Then parents in occupation 1 must be indifferent between continuing with occupation 1 and shifting their children (after re-optimizing fertility) to occupation 0. Denote by $\hat{n} = n_1(0)$ the number of children that an occupation-1 parent would choose if she were switching her progeny to occupation 0. Then, by indifference, we have

$$u(c_1) + \frac{\delta n_1^\theta u(c_1)}{1 - \delta n_1^\theta} = u(w_1 - r_0(w_1)\hat{n}) + \frac{\delta \hat{n}^\theta u(c_0)}{1 - \delta n_0^\theta} = u(c_1) + \frac{\delta \hat{n}^\theta u(c_0)}{1 - \delta n_0^\theta},$$

where the second equality follows Lemma 1: total expenditure on children must be equalized at a switch point. Consequently, continuation utilities are equalized, and using the constant-elasticity specification, we obtain

$$\left(\frac{c_1}{c_0}\right)^{1-\rho} = \frac{u(c_1)}{u(c_0)} = \frac{\hat{n}^\theta(1 - \delta n_1^\theta)}{n_1^\theta(1 - \delta n_0^\theta)} = \frac{r_1(w_1)^\theta(1 - \delta n_1^\theta)}{r_0(w_1)^\theta(1 - \delta n_0^\theta)}.$$

Combining this equation with (A13), we see that

$$\frac{r_1(w_1)n_0^{\theta-1}(1 - \delta n_1^\theta)}{r_0(w_0)n_1^{\theta-1}(1 - \delta n_0^\theta)} = \frac{r_1(w_1)^{\theta/(1-\rho)}(1 - \delta n_1^\theta)^{1/(1-\rho)}}{r_0(w_1)^{\theta/(1-\rho)}(1 - \delta n_0^\theta)^{1/(1-\rho)}},$$

or

$$\text{(A15)} \qquad \frac{r_1(w_1)r_0(w_1)^{\theta/(1-\rho)}}{r_0(w_0)r_1(w_1)^{\theta/(1-\rho)}} = \frac{n_0^{1-\theta}(1 - \delta n_1^\theta)^{\rho/(1-\rho)}}{n_1^{1-\theta}(1 - \delta n_0^\theta)^{\rho/(1-\rho)}}.$$

Because our steady state has downward mobility, we have $n_1 > n_0$, which implies that the right-hand side of (A15) is strictly smaller than 1 under condition (a).

On the other hand, the left hand side is given by

$$\frac{r_1(w_1)r_0(w_1)^{\theta/(1-\rho)}}{r_0(w_0)r_1(w_1)^{\theta/(1-\rho)}} = \frac{r_1(w_1)^{(1-\theta-\rho)/(1-\rho)}r_0(w_1)^{\theta/(1-\rho)}}{r_0(w_0)} \geq 1,$$

given the assumptions of the Proposition as well as R.1 and R.2. This contradiction completes the proof.

PROOF OF PROPOSITION 6. Recall the characterization of steady states in Lemma 2: a steady state with downward mobility satisfies $w^*(\lambda) = w_1(\lambda)$ and with upward mobility satisfies $w^*(\lambda) = w_0(\lambda)$. So for either of these kinds of steady states, it suffices to show (via standard transversality arguments, e.g., see Mas-Colell, Whinston and Green (1995, Proposition 17.D.3)) that an increase in educational cost parameter ν causes $w^*(\lambda)$ to strictly increase, for any λ.

Note initially that at any given λ, V_0 is unchanged, while V_1 must fall as ν rises. This follows from the fact that in steady state it is optimal for unskilled parents to not educate their children, so they must be unaffected by the rise in ν. And it is optimal for skilled people to educate their children, so they must be worse off when ν rises.

Next, manipulate the first order condition (2) for fertility decisions for occupation i to obtain the following equivalent version:

$$r_i^{\theta} u'(w_i - E_i)E_i = \delta E_i^{\theta}.\theta V_i \tag{A16}$$

where $E_i \equiv r_i(w_i)n(w_i, i)$ denotes total expenditure on children who are trained for the same occupation, and all variables are evaluated at the given skill ratio. It is evident that all decisions of unskilled parents are unaffected.

Consider first the case where θ is positive. Then θV_1 falls as ν rises. So (2) implies that fertility n_1 of skilled parents must fall (where $n_i \equiv n(w_i, i)$). Now observe that (A16) can also be written as

$$u'(w_i - E_i)E_i = \delta n_i^{\theta}.\theta V_i \tag{A17}$$

Since n_1 and θV_1 both fall, it follows from (A17) that E_1 must fall. Since $\theta > 0$, and E_0 is unaffected, it follows that parents are less inclined to educate their children (as they tend to maximize child expenditures), and w^* must rise.

Now suppose $\theta < 0$. In this case θV_1 rises as ν rises. Equation (A16) now implies that E_1 rises. Since parents make human capital decisions for their children on the basis of minimization of total expenditures, they are again less inclined to educate their children, and w^* must rise.

Since the wage functions $w_i(\lambda)$ are unaffected by the change in ν, the result now follows for steady states with either upward or downward mobility. Steady states with zero mobility must satisfy the condition that $n(w_1(\lambda), 1) - n(w_0(\lambda), 0) = 0$, and an increase in ν must cause the left-hand-side of this equation to fall when θ

is positive (n_1 must fall while n_0 is unaffected, from (2)).

PROOF OF PROPOSITION 7. Given expression (14) for w^*, it suffices to show that the derivative of $\frac{V_1(\lambda)}{V_0(\lambda)}$ with respect to f and x are respectively positive and negative in the case where $\theta > 0$ (and signs reversed in the case that $\theta < 0$), at any steady state with upward mobility.

Recall that

$$V_i(\lambda) = \max_{n_i}[u(w_i(\lambda) - (\psi w_i(\lambda) + f + xi)n_i) + \delta n_i^\theta V_i(\lambda)]. \tag{A18}$$

Applying the Envelope Theorem to this optimization problem we have

$$\frac{\partial V_i(\lambda)}{\partial f} = -\frac{u'(c_i(\lambda))n_i(\lambda)}{1 - \delta n_i(\lambda)^\theta} \tag{A19}$$

where $n_i(\lambda)$ denotes the optimal choice of n_i, and $c_i(\lambda) \equiv w_i(\lambda) - (\psi w_i(\lambda) + f + xi)n_i(\lambda)$. Therefore

(A20)

$$V_0(\lambda)\frac{\partial V_1(\lambda))}{\partial f} - V_1(\lambda)\frac{\partial V_0(\lambda))}{\partial f} = -V_0(\lambda)\frac{u'(c_1(\lambda))n_1(\lambda)}{1 - \delta n_1(\lambda)^\theta} + V_1(\lambda)\frac{u'(c_0(\lambda))n_0(\lambda)}{1 - \delta n_0(\lambda)^\theta}.$$

Next, note that in any steady state with upward mobility we have $n_1 < n_0$, which in turn implies that $c_1 > c_0$ (in order to ensure that $V_1 > V_0$).

In the case of positive utility where $\theta > 0$, it follows that expression (A20) is positive.

In the case of negative utility where $\theta < 0$, use expression (A18) for the value function to see that expression (A20) reduces to

$$\frac{u(c_1)}{1 - \delta n_1^\theta}\frac{u'(c_0)n_0}{1 - \delta n_0^\theta} - \frac{u(c_0)}{1 - \delta n_0^\theta}\frac{u'(c_1)n_1}{1 - \delta n_1^\theta} \tag{A21}$$

Since $n_0 > n_1$, it suffices that $[-u(c_1)].u'(c_0) \geq [-u(c_0)].u'(c_1)$ for (A21) to be negative. This is ensured by the property that $\frac{u'}{-u}$ is non-increasing.

It follows that $\frac{\partial w^*}{\partial f} < 0$.

A simpler argument ensures that $\frac{\partial w^*}{\partial x} > 0$. The argument is simpler because the V_0 function is locally unaffected by a small rise in x, while $\frac{\partial V_1}{\partial x} = -\frac{u'(c_1)n_1}{1-\delta n_1^\theta} < 0$.

Hence the w^* curve shifts down following an increase in f or a decrease in x, while the position of the w_0 curve is not affected. If the steady state is locally stable, a small rise in f or fall in x must cause the skill ratio to rise.