Schultz, T. Paul

Working Paper
The Fertility Transition: Economic Explanations

Center Discussion Paper, No. 833

Provided in Cooperation with:
Economic Growth Center (EGC), Yale University

Suggested Citation: Schultz, T. Paul (2001) : The Fertility Transition: Economic Explanations, Center Discussion Paper, No. 833, Yale University, Economic Growth Center, New Haven, CT

This Version is available at:
http://hdl.handle.net/10419/98308

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Abstract

Economic explanations for the fertility transition focus on the role of returns to schooling, especially for women, which have encouraged women to obtain more education and facilitated the rise in women’s wages relative to men’s. The private opportunity costs of children have therefore increased, and parents have been motivated to substitute child schooling for additional births. Declines in fertility have proceeded unevenly, first across the high income countries, and more recently across the low income countries. The cross-sectional differentials in fertility are also frequently analyzed in household surveys, suggesting parallels with the cross-country comparisons. At an aggregate level, states have simultaneously legislated socialized support for the consumption of the elderly, which has eroded the incentives for childbearing, and subsidized child human capital through schools and public health programs, which has encouraged parents to demand fewer, higher quality, children.

Key words: Fertility Transition, Women’s Schooling, Women’s Wages, Child Mortality

JEL Codes: D19, J10, J13, N30
1. Approaches of Economics and Demography to the Fertility Transition

With the establishment of modern economic growth, fertility has tended to decline, first across high income countries starting largely in the last quarter of the nineteenth century, and then occurring even more rapidly across most low income countries in the last quarter of the twentieth century (Kuznets, 1966; Schultz, 1981). Economic models of the allocation of time and wealth of families among production and consumption activities suggest several hypotheses linking changes in the economic environment in which people live to their reproductive demands and fertility (Becker, 1960; Mincer, 1963). Empirical counterparts for the price, income, and environmental constraints on people’s fertility choices have been used to explain aspects of fertility and other related outcomes of the family, strengthening the view that this household demand model is a source of insights into the determinants of the historic fertility transition (T.W. Schultz, 1974; Becker, 1981).

This economic approach to fertility determinants has been criticized by those who believe that fertility is not a rational response to changing economic conditions (Cleland and Wilson, 1987). The proposed alternative to the economic approach, called “ideational change” emphasizes social interactions, but has not yet evolved to a form that is empirically distinguishable. However, research is trying to quantify the role of groups external to the family, such as peers and neighborhoods, and to interactions between family members as postulated in cooperative and non-cooperative game strategies within families. But because these groups are generally self selected, it is complex to specify the factors that affect the formation of the group and do not also affect the behavior of individuals who enter such groups (Heckman, 1979; Manski, 1995). There is active debate on how to apply the economic framework and to exploit statistically cross sectional and time series variation in fertility to improve our understanding of the determinants of fertility (e.g., Schultz, 1997).
There is a fundamental methodological divide between those who want to describe a “structural model” that involves the interaction between choices made at the household level, including those choices spanning an adult’s lifetime, such as between a woman’s fertility, marriage, and labor force history, and those who investigate a “reduced form model” where only the fixed constraints on household choice are considered as potential explanatory variables -- market prices, income sources, technologies, and the environment. It is widely recognized that the associations between lifetime choices represent more (or less) than a simple causal effect of one choice on the other. These covariations between family choices embody possible effects transmitted in both directions, effects of unobserved factors which impact both outcomes, and heterogeneity due to differences in individual preferences (Schultz, 1981). Strong identification restrictions are generally required to estimate from non-experimental data the critical parameters to such “structural models,” and therefore the empirical conclusions drawn from them are often controversial, because they depend on a host of untested and debatable modeling choices (Manski, 1995).

The alternative approach is to estimate fertility as a function of explanatory variables that are not family choice variables or not likely to be impacted by unobserved factors that could also affect fertility. By delimiting the potential list of explanatory variables, these “reduced form” relationships represent a summation of the direct effect of the environmentally fixed constraints on the couple’s fertility, plus the possible indirect effect of the same constraints on fertility that operate through the modification of other lifetime choice variables, such as the woman’s age at marriage, extent of labor force activity outside of her home, contraceptive and sexual behavior, etc. The limitation of reduced form estimates is that they do not necessarily clarify the mechanism by which the constraints impact fertility, for example through the timing of marriage, or the adoption within marriage of various so-
called “natural” or “unnatural” forms of birth control. The strength of reduced form estimates is that they claim to provide an unbiased approximation for the total effect of the constraint on fertility, which is often a central concern of the social scientist, forecaster, or policy maker.

Some historical demographers assign special importance to the moment when married women begin to restrict their birth rates in response to their prior number of births, or surviving children (Coale and Watkins, 1986). Although this may be a well-defined and interesting threshold that predicts the onset of the long term decline in fertility in certain European populations, other populations may adopt other mechanisms to reach their changing fertility goals, such as a change in median age at marriage (i.e. Malthus’ preventive check), an increase in celibacy or the proportion who never marry in a birth cohort (Weir, 1984), an increase in the spacing of births within marriage, as appears to be evident in Africa today (Caldwell et al., 1992; Shapiro and Tabashe, 2000), or a change in non-marital fertility. Even in the region of Northern Europe, standard measures of the emergence of parity-specific fertility control within marriages are not precisely defined by existing data (Guinnane, et al. 1994). Thus, economists generally consider comprehensive measures of cohort fertility, and do not focus on one component, such as marital fertility.

2. Stages of the Demographic Transition

The “demographic transition” refers to a sequence of three periods. In the first period, fertility and mortality rates are high and mortality tends to be highly variable, with population growth fluctuating widely, about a moderate long run trend of growth or decline. Short run responses in fertility (and mortality) can be attributed to cycles in weather or harvests or possibly to other exogenous variables using time series statistical methods, such as vector-auto-regression.

In the second period, age-specific mortality rates decrease gradually, raising life expectancy
at birth from 30-35 years in pre-industrial Europe, to 70-75 today in the high income countries, and from 25-30 years in the low income countries in the 1920s, to levels which today range from about 45-73, excluding countries ravaged by war or the AIDS epidemic. This second period is one of accelerating population growth, which some observers such as Malthus attributed to improved food supplies, and more specifically to nutrition, often initiated by technical progress in agriculture, followed by improved industrial productivity, and advances in transportation and communication (Fogel 1999). A free press capable of publicizing food shortfalls in the twentieth century may have further alleviated excess mortality in the wake of periodic famines. Improvements in private and public health technologies are assigned an important role in reducing mortality, but not much before the start of the twentieth century.

In the third period, the secular decline in fertility begins, after which the number of births per woman falls by more than half, from five or six to about two in the high income countries, and six to eight to about two or three in most low income countries. When historically high fertility levels have declined by more than 10 percent, they are not expected to rise again on a sustained basis (National Research Council 2000). Medium term swings in fertility are nonetheless distinguished from time series, occurring in response to business cycles (e.g., the post-World War II baby boom), wars, and economic shocks, as during the transition from centrally planned to market oriented economies in the 1990s in Eastern Europe and Russia, or recently due to economic reversals in Sub-Sahara Africa.

3. Economic Determinants of Fertility

Notestein (1945) described the early dimensions of the demographic transition, and speculated on the economic pressures which he expected to reduce fertility in the low income world: rising cost of children in urban life, individualistic promotion of child health and education. Economic models
of a lifetime fertility choice emphasize different facets of this decision process, but the trade-off between quantity of children and their quality (i.e., education and health) has become a central feature (Becker 1960, Mincer 1963, Schultz 1981). Because women’s time in child-care is a large fraction of the modern opportunity cost of fertility, her allocation of time is treated as jointly determined with fertility. Many cross-sectional and a few time-series studies have proceeded to relate statistically these fixed constraints to fertility outcomes.

A couple’s lifetime wealth (endowments of human capital and nonhuman capital) is expected to increase with modern economic growth, and to encourage parents to demand more children, if the relative costs and functions of children did not change (Becker 1960, Schultz 1981). However, for reasons discussed below, the opportunity cost of children increased along with the cost of other goods and activities that complement the rearing of children, whereas the private price of some substitutes for children has declined, such as old-age pensions and health care, or have been socialized and subsidized by governments. Marked improvements in the technology of birth control were developed in the 1950s translating into the rapid diffusion of the pill (oral steroid contraceptive) and the IUD (intrauterine device) to most regions of the world. The monetary and psychic cost and inconvenience (independent of coitus) of avoiding too many births was reduced, and should thereby have reduced the number of births. Family planning programs are not, however, a necessary condition for the occurrence of a fertility transition, because birth rates declined in high income countries without the benefit of new control technologies, or of family planning extension activities designed to foster the adoption of birth control. Finally, improvements in child health technology increased the proportion of births that survived to adulthood. These health developments put downward pressure on birth rates, under two assumptions: Parent reproductive goals are defined
primarily in terms of their number of surviving children, and these demands for survivors are highly price inelastic (Schultz 1981).

The actual expenditures of a couple on each of their children is not a suitable indicator of the “market price” of a child on which a couple would base their demand for children. Expenditures per child are also a family choice, which Becker called the demand for “child quality.” The market-determined prices of the inputs that households combine to rear children are the relevant prices that condition fertility. But, there is no consensus on which inputs to children are most important and how their unit prices are measured?

If one accepts Becker’s (1960) hypothesis that quality and quantity of children are viewed by parents as substitutes, and much of the variation in child quality is embodied in the education parents give their child, then the net price of schooling would affect negatively the demand for quality and affect positively the demand for quantity of children. There is growing evidence in the twentieth century that wage returns to schooling increased, and this can be viewed as a decline in the price of schooling, encouraging parents and societies to invest more resources in the education of children. If parents view schooling per child as a substitute for numbers of children, as is now often empirically confirmed (e.g., Rosenzweig and Wolpin 1980), the increasing wage return to schooling among workers should encourage parents to substitute child schooling for number of children. The increase in wage returns to schooling can be attributed to a technological bias favoring skilled labor, or to a complementarity in production between the schooling of workers and the accumulating stock of technical knowledge available in the global economy.

Becker (1981) also assumed that the income elasticity on the demand for child quality was a larger positive magnitude than the income elasticity on the demand for child quantity. This might
also have encouraged parents to substitute child quality for child quantity as their incomes increased from one generation to the next. Here was a second possible explanation for the anomalous fall in fertility in a period when *per capita* income was increasing.

4. The Increasing Value of Parents’ Time as an Opportunity Cost of Children

In the twentieth century, the fraction of national income paid out to labor has risen, despite the increase in capital employed per worker and the decrease in hours worked per year per worker (Kuznets 1966, Fogel 1999). Wage rates have grown relative to *per capita* income, implying that consumption activities that use more time of the consumer have become relatively more expensive over time. Children are thought to be a more “time-intensive commodity” than the average of all consumption commodities, and thus the opportunity costs for households to produce and consume a child have risen as parent wages have risen, offsetting some part of the increased demand for children that might have been expected in a period when *per capita* income is rising (Mincer 1963, Becker, 1965).

The household production-consumption model drew attention to gender specialization in social production, and to the “female time-intensity” of children, thereby suggesting the woman’s value of time could be a key component in the price of children. In environments where the opportunity cost of female (i.e., mother’s or child-care giver’s) time is higher, fertility is expected to be lower, other things equal (Schultz 1981). The market wage of women might thus be an “opportunity price” variable, and it has notably increased in many countries in the twentieth century at a time when fertility has fallen. But just as “parental expenditures per child” does not represent a fixed market-determined price of a child, the woman’s market wage does not represent a fixed price of a woman’s time input to child rearing. First, the wage a woman receives in the labor force is
affected by her prior job experience, which accumulates at different rates depending on her preferences among types of work and the allocation of her time. Consequently, a woman’s wage is an endogenous choice variable closely linked to her lifetime allocation of time, which in turn is intimately related to her fertility decisions (Mincer 1963).

A second problem with the woman’s wage variable is that it is observed only for a self-selected fraction of women who work for a wage in the observation period, and this fraction has tended to rise over time in many populations, and to differ by socioeconomic strata within a population. It is possible that women who work for wages are those who receive an unusually high market wage offer, suggesting that the observed average wage of women overstates the value of time of an average woman. Conversely, women who are unusually productive in non-wage work, and thus are less likely to participate in the wage labor force, may still have an above average market-wage offer, in which case the observed average wage of women would understate the value of time of an average woman. Sample selection bias could distort the wage as a measure of the value of women’s time, and thus bias estimates of the woman’s wage effect on fertility, probably overstating the causal connection in time series (Heckman 1979, Schultz 1981).

5. **Determinants of the Value of Women’s Time and the Cost of Children**

The variation in the opportunity price of children that arises from the variation in women’s opportunity value of their time remains a key economic determinant of fertility, and econometric procedures are required to estimate the importance of this connection for the fertility transition. In other words, because the observed wage of women cannot be treated as a fixed determinant of fertility, it cannot be included in a reduced form equation for fertility. How then, does one approximate in a more structural model the effect of employment and productive opportunities for
women on the opportunity cost of a woman’s time, which in turn affects fertility?

One approach to this problem is to employ the woman’s education as a proxy for her wage; schooling tends to be the best variable available for predicting in the cross-section women’s wages, just as it is for men’s wages. For most women their human capital investment in schooling is completed before childbearing starts, and thus the schooling decision can be viewed as predetermined with respect to fertility. If the schooling decision is made by the child’s parents, it may also be argued that the woman’s schooling is unaffected by her adult preferences between labor market and fertility choices, and then schooling may be a satisfactory “instrumental variable” for predicting the adult lifetime wage opportunities for women, which constrain the choice of fertility.

A second approach is to treat the structure of wages in the local labor market as fixed and thus independent of the woman’s preferences. This appears to be a reasonable strategy, proceeding from the individual wage to the community averaged wage, and thus eliminating the idiosyncratic individual variation in preferences between working and fertility. Analyses of fertility can then include all women based on the regional wage, and not be restricted to a selected sample of women who are currently reported working for a wage. But another problem could still bias the estimated effect of women’s regional wage opportunities on regional or individual fertility. Because an increased in the labor supplied by women in a region will tend to depress the market-clearing wage women receive locally, any omitted regional variables that contribute to greater female labor force participation are expected to also be correlated with lower wages, and hence be correlated with higher fertility. Consequently, these regional omitted labor supply variables would introduce a positive bias in the estimated effect of woman’s regional wage on fertility, masking the theoretically expected negative impact of aggregate demand-induced variation in women’s wages on fertility.
One estimation strategy to deal with this regional omitted variable problem is to specify a regional variable that shifts only the aggregate regional economy’s demand for female labor (relative to the demand for male labor), and thereby affects the local wage available to women (relative to men). Variables of this form which shift the aggregate demand schedule for women’s labor are candidates for being an “instrumental variable” that can account for demand-induced variation in women wages, and can thus be used to obtain an unbiased estimate of the effect of women’s wages on fertility. One such demand variable is the regional location of natural resources or the composition of industries, which influence the gender composition of employment opportunities and wages.

Coal mines, for example, employ mostly men, and communities dependent on coal mines for their livelihood have relatively low wages for women. Historically, these coal communities in England and the United States reported relatively high fertility levels, other things equal (Haines 1979). In the Northern provinces of Sweden where the price and quantity of timber increased at the end of the nineteenth century due to exports, men’s wages increased relative to women’s wages, and fertility remained high in these regions well into the twentieth century. By contrast, fertility in the Southern provinces of Sweden declined after the 1860s, as women’s wages increased relative to those of men, in response to higher output prices for exports of livestock, dairy and food manufactured products, which employed relatively more women (Schultz 1985). Natural resource exports strengthen government finances by providing an inelastic source of taxation, notably in the case of oil exports, with a tendency to sustain customary social institutions that have been associated with relatively low levels of female education, low female labor force participation, and high levels of fertility, holding constant for other factors (Schultz, 1994).

Labor economists have noted since the 1960s that in regions where the share of employment
in certain industries, for examples the services, is relatively large, female labor force participation rates tend to be higher. The importance of part-time, less-skilled, and physically less-demanding jobs may thus be associated with a larger share of women in the labor force, and a rise in the economic status of women relative to men. This form of “female bias” in the industrial composition of local employment is found to predict increased women’s wages relative to men’s and lower fertility.

Changes in fertility over time can then be explained by changes in factors external to the household sector which govern the expansion of employment opportunities in industries that hire a disproportionate share of women (or men). The price of the industry’s output when it is determined in international markets is one such factor. Terms of trade indexes linked to gender-specific employment and wages opportunities can predict women’s wages relative to men’s and provide a basis for estimating how demand-driven variation in women’s employment opportunities influence fertility. However, these changes in relative prices and employment prospects must be perceived by couples as sufficiently long term in nature that they feel the need to adapt accordingly their reproductive goals.

6. Evidence from European Fertility Transitions

The decline in the price of food grains in Europe in the nineteenth century led to such a profound reallocation of labor and allied resources. With the opening of the Erie Canal in 1825 and the reduction in ocean transportation costs, exports from the USA, Russia and other world grain exporters depressed grain markets across Europe. Although restrictions of trade attempted to protect domestic agriculture, European farmers gradually adapted to their changing comparative advantage and left agriculture for urban jobs or emigrated, whereas those who remained in agriculture reallocated their effort from grain to livestock production and to the expansion of food and dairy
processing industries. Because grain production traditionally employed more male labor, and livestock and dairy production used more female labor, these changing output prices boosted women’s wages relative to men’s. This reallocation of resources occurred more rapidly where economies were more open to international trade in agricultural commodities. This transformation of composition of European agriculture or its consequences on the productive roles of women and men is not quantitatively studied in the Princeton European Fertility Project, although these changes could have influenced the diffusion of the fertility transition (Coale and Watkins 1986).

In Sweden, as noted earlier, these agricultural relative price changes were closely associated with the improvement in women’s wages relative to men’s, and with declines in fertility, by province over time, starting after 1860. The level of real wages for men was related to younger childbearing, but not to a change in total fertility rates. About one-third of the 25 percent decline in total fertility rates for Sweden in the period 1860 to 1910 can be attributed to these externally driven increases of women’s wage opportunities, whereas the rest of the decline in Swedish fertility in this half century is associated with urbanization and the increase in child survival, under the assumption that these are exogenous developments (Schultz 1985). A similar demographic pattern unfolded in Denmark, which was also relatively open to international agricultural trade, while its demographic transition was only slightly less rapid than in Sweden.

The German fertility transition has also been analyzed from 1880 to 1910. When regional persistent or fixed-effects are introduced in the analysis of times series of cross sections, the evidence of strong partial effects of industrialization and urbanization emerge (Richards 1977). Smaller regional units were examined in Prussia from 1875 to 1910, with the same regional fixed-effect specification for fertility (Galloway et al 1994). However, the female labor force participation rates
are used to explain fertility in what was otherwise a reduced form specification, as though the changing labor force participation of women was a “fixed constraint” on fertility decisions. Not surprisingly, this choice variable reflecting women’s time allocation accounts for the largest fraction of the fertility decline, and if it is jointly determined by women with their fertility, as seems likely, the reported estimates are subject to simultaneous equation bias and are difficult to interpret as evidence for any causal model.

7. Advancement in Women’s Schooling

Another widespread development in the world, which has contributed to increasing the cost of children relative to most other forms of consumption, is the advance in women’s schooling. In the USA, by the start of the twentieth century, the gender gap in schooling was modest, except for college education. But this gap remained substantial in most countries in Southern and Eastern Europe until after World War I. Moreover, women’s schooling attainment began to gain relative to that of men in low income countries only in the last few decades. This progress has been particularly sluggish in South Asia where the fertility transition has also been most gradual. Given the empirical tendency for wages to increase for women by at least the same percentage as they do for men when they obtain an additional year of schooling, the closure in the gender gap in schooling implies that women’s wages are likely to continue to increase in most parts of the low income world for some time, and indeed increase at a more rapid rate than will the wages of men (Schultz 1995). The factors that have triggered these differential trends in the schooling of women and men have not been thoroughly studied, but the expansion of women’s schooling closely parallels the decline in fertility across high and low income countries since 1960. Indeed, changes in women’s schooling accounts for the largest share of the changes in total fertility rates observed in low income countries since 1960,
and in Africa since 1980 (Schultz 1994). Even in an environment where economic progress has been minimal since the 1960s – Kinshasa City in Congo – the increasing rate of female education is the key factor explaining different levels of fertility across the population, and accounting for the decline in overall fertility in the 1980s and 1990s (Shapiro and Tambashe 2000).

However, educated women could have lower fertility for many reasons, in addition to the greater opportunity cost of their time in child rearing. The social and intellectual advantages that educated women enjoy help them in deciphering, adopting and using effectively, new and old forms of birth control, and thereby avoiding unwanted births.

Female schooling may also affect fertility through its impact on the risks of child mortality. Censuses and surveys after World War II have documented for dozens of countries that child mortality rates tend to be 5 to 10 percent lower for mothers who have completed an additional year of schooling (Schultz 1981, 1997). However, the proportionate decline in fertility associated with women’s education is larger than the proportionate increase in their child survival rates. Therefore, better educated women have smaller numbers of surviving children, which implies they contribute less to the overall rate of population growth than their less educated peers (Schultz 1994).

Education of women is also closely related to their labor force participation outside of their family, holding constant for their husband’s income or education (Schultz 1981). Household models of production and consumption generally assume that family labor supply choices are jointly determined, and women’s education is expected to affect not only lifetime fertility, but also the timing of marriage and the matching of the productive characteristics between husband and wife in the marriage “market.” Within such a framework of family coordinated decision making, it is clearly difficult to assess how one of these jointly determined life-cycle choices made by women affects
another of their choices, as stressed earlier. To decompose the variation in fertility into that which is due to the timing of marriage, marital fertility rates, and non-marital fertility rates is a suggestive arithmetic exercise, but one that can misguide statistical analysis because the three components are likely to be stochastically interdependent. Consequently, when analyses of fertility differentials during the demographic transition focus only on marital fertility differentials, correlations can be causally understood only if marriage behavior and fertility demands are behaviorally independent of each other, a working assumption that few social scientists would adopt (Coale and Watkins 1986, Cleland and Wilson 1987).

8. Conclusions

Economic explanations for the fertility transition recognize the catalytic contribution of increasing returns to schooling, especially for women, which has encouraged women’s education relative to men’s, and fostered the rise in women’s wages. These attributes of development are not equally evident in all countries. Declines in fertility have therefore proceeded quite unevenly, first across the low income countries, but following a pattern of economic developments that has been widely quantified in household surveys and suggestively followed in cross-country comparisons. At an aggregate level, states have simultaneously legislated socialized support systems to provide for the consumption requirements of the elderly, and for the school and health investment requirements of the young, which may have also on balance, reduced the private demands of parents for children.
Bibliography


