1977

Experience with Pregnancy, the Demand for Prenatal Care and the Production of Surviving Infants

Eugene M. Lewit
Graduate Center, City University of New York

77-13,855

LEWIT, Eugene Myron, 1946-
EXPERIENCE WITH PREGNANCY, THE DEMAND FOR PRENATAL CARE AND THE PRODUCTION OF SURVIVING INFANTS.

City University of New York, Ph.D., 1977
Health Sciences, human development

Xerox University Microfilms, Ann Arbor, Michigan 48106

EXPERIENCE WITH PREGNANCY, THE DEMAND FOR PRENATAL CARE AND THE PRODUCTION OF SURVIVING INFANTS

by

EUGENE M. LEWIT

A dissertation submitted to the Graduate Faculty of Economics in partial fullfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1977

This manuscript has been read and accepted for the Graduate Faculty in Economics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

date — Chairman of Examining Committee

date — Executive Officer

Michael Grossman

Harold Hochman

Elliot Zupnick

Supervisory Committee

The City University of New York

Abstract

Experience with Pregnancy, the Demand for Prenatal Care
and the Production of Surviving Infants

by

Eugene M. Lewit

Adviser: Professor Michael Grossman

The object of this research is to develop a model of household demand for prenatal care and attempt to measure the productive value of prenatal care per se on infant health as measured by survival. Traditionally, infant mortality rates have been used as indices of a nation's health status. Since the U.S. has lagged significantly behind other developed nations in reducing infant mortality since the mid-1950's, there have been charges of a malfunction in the U.S. health delivery system. Particularly in the area of infant health, critics have charged that more prenatal care inputs are needed and that they should be directed specifically towards so-called high risk mothers. Others have questioned the value of input intensive prenatal care, claiming its marginal product is low, cost high and efficacy unproven.

An economic model is developed in which the demand for healthy children is viewed as being derived from the demand for children per se. In a world where families cannot substantially effect the outcome of individual pregnancies by varying inputs, it is demonstrated that measured infant mortality rates will be not only a function of health status but

also fertility decisions. In a world where families can vary inputs, it is argued that prenatal care, as the most pregnancy relevant related input, should be a good index of the total demand for pregnancy related inputs. It is argued that the level of inputs will be positively correlated with income, tend to increase in families who have experienced pregnancy losses and decrease as family size increases, particularly if marginal children are less "wanted" as family size increases.

Demand and production relationships are estimated using data from the 1970 New York City birth cohort. The data set consists primarily of birth and linked death certificates for the period January to June, 1970 and contains 54,000 observations after editing.

Several different dependent variables are utilized to estimate the demand for care. They include a dichotomous care/no care variable, the interval to the first visit and the number of visits. Significant empirical findings include: (1) the decision whether or not to seek care is most strongly influenced by legitimacy status; (2) the demand for care is effected by past experience as predicted by the model in that families with more live children demand less care and those with a history of losses demand more care; (3) substantially less care is demanded by blacks, foreign born and Puerto Rican born mothers even when other variables are accounted for; (4) less care, other things equal, is obtained in specially designated Maternal and Infant Care Project areas, despite the presence of these special projects to encourage the use of care by high risk mothers of low socio-economic status; (5) the amount of care a mother receives is substantially determined by obstetrical protocol and does not seem to reflect her previous pregnancy experience.

Outcome measures include birth weight, infant death, neonatal and postneonatal death. Regarding birth weight significant findings include: (1) birth weight differentials attributable to race, ethnicity, nativity or legitimacy characteristics are substantially reduced by taking account of differentials in the level of care received; (2) the net gain in birth weight attributable to a full complement of prenatal care (303 grams) as compared with no care is substantial when comapred with the birth weights of "high risk" infants; (3) previous experience of pregnancy successes and losses are reflected by increments or decrements in birth weight.

In comparing results of outcome regressions for neonatal and post-neonatal mortality, it is found that other things equal, prenatal care has a positive effect on survival during the neonatal period but no effect during the postneonatal period. Hence, it is argued that care per se has real value in improving pregnancy outcomes and is not primarily acting as a proxy variable for "wantedness" or other unmeasured inputs.

The results of using FIML logit estimators on the dichotomous dependent variables biased on a subsample of observations do not agree with the OLS estimates based on the entire sample. It is suggested that econometricians need to more fully explore the relationship between the value of these two techniques, particularly in very large data sets.

Acknowledgements

As this dissertation had an extremely long gestation period many individuals had a hand in its birth. My greatest debt is to Professor Michael Grossman who served me not only as an adviser but as a friend, particularly when the going got rough. He was always willing to read drafts carefully and critically and was always available to discuss a troublesome point. He shouldered the burden of seeing the thesis through to its conclusion when substantial personnel changes left him with an increased work load and me without my previous advisers. He coaxed me to finish with a carrot, not a stick and celebrated with me when the job was done. For these and other reasons I am very greatful.

I owe a substantial debt to Professor Victor Fuchs who first introduced me to the study of health economics. He supervised the initial stages of this research and provided many incisive and helpful comments. He also facilitated my working as a junior member of the health economics project at the National Bureau of Economic Research where much of this work began and where I had the opportunity of interacting with many others interested not only in child health but also the economics of fertility decision making, a closely related area.

Although I am not clear as to the extent of the NBER commitment to this study, I do know that when I was there I was encouraged to pursue this research and facilities and personnel were available to me to help me when my other responsibilities left me too little time to move this project along as I might have liked. Among the fine group of NBER research assistants whose time I was able to claim were John Wolfe, Christy Wilson, Don Wright and Terry Eschavez. At the time of my sojourn at the NBER, the health economics project was supported by the Robert Wood Johnson Foundation and the National Center for Health Services Research.

As my work entered its final stages, Professors Elliot Zupnick and Harold Hochman graciously agreed to serve on my defense committee and provided many excellent suggestions as to how to make the study more intellegible to the nonspecialist. Earlier I benefited from readings and comments by Barry Chiswick, Fred Goldman, Robert Willis, Melvin Reder, and Finis Welch. I've received helpful comments on this research at presentations at the National Bureau of Economic Research, CUNY Graduate Center and Yale University. Ann Williams was particularly helpful in serving as the discussant of a presentation of some of my results at a HERO session in September, 1976.

At Mount Sinai, Rajneesh Ghei and Barbara Chamberlain, as research assistant and secretary respectively, have performed heroically in trying with me to meet almost impossible deadlines. My friend and colleague, Edward F.X. Hughes, has been most considerate and understanding when my work on this study took my time and our staff time away from other projects.

It is traditional to close the acknowledgements section by thanking one's spouse and let me not deviate from this tradition. Like most wives, Judy suffered long periods of neglect when the thesis seemed to come before my responsibilities to her and our children. She also filled in as a research assistant usually working late hours just before I was scheduled to present a seminar somewhere. Through my first hand

observations of, participation in and discussions with her about the birth of our own two children, I gained information about prenatal care procedures and the pregnancy process. But most of all, I appreciate the unspoken way in which she seemed to be able to encourage me when the going was rough and to let me know of her pleasure when things went well - it was these little things that made it all seem so much more worthwhile.

Table of Contents

List of Tables

Chapter VI:

Chapter VIII:

List of Figures

Chapter I - Introduction

Infant mortality has long been an area of interest to students of health. Variations in infant mortality rates and the determination of factors associated with favorable pregnancy outcomes have been of substantial concern because the high vulnerability of children during the first year of life would tend to argue that the implementation of policies designed to reduce infant loss could thereby significantly raise health levels. Moreover, whether in cross sectional or time series studies, the reduction in infant mortality rates has tended to reflect favorable improvements in living, economic, and health conditions.

Economists and demongraphers have been interested in infant mortality rates and the interaction of different levels or expectations of infant mortality with fertility. The result of such interactions may importantly determine population growth. In this regard, a traditional concern and one that still exists in developing countries is that excess population growth will strain the other resources of the economy and lead to a relative or absolute decline in output per capita because of the scarcity of other inputs.

While the area of the interaction between population growth and economic growth has been a traditional area of interaction between demographers and economists, recently economists have turned their attention to less traditional topics of concern. New perspectives have begun to be developed for a more comprehensive economic approach to the study of infant mortality rates and the continued responsiveness of infant mortality rates to aggregate economic growth has been questioned (Fuchs, 1974). Attention has focused on the workings of health care system in the United States.

Critics have pointed to the rapid growth of expenditures on health, rapidly increasing prices of the health care, and the increasing share of GNP devoted to health and have voiced concern about whether the benefits from such an increased utilization of resources warranted this level of expenditures.

One of the indices that has been of particular concern in this discussion has been the high rate of infant mortality in the U.S. Despite the very high levels of expenditure on health, the U.S. has lagged significantly behind other developed countries in the reduction of infant mortality, so much so that during the 20-year period from 1950 through 1970, the United States' infant mortality rate dropped from a position of seventh lowest rate among 15 industrialized countries to highest (Chase, 1972). Over that twenty-year period, the U.S. experienced a decline in the rate of infant mortality significantly lower than almost any other developed nation. As if this were not adequate reason for concern on the part of critics, it has been pointed out that the United States uses a relatively expensive technology in the production of healthy infants as compared to certain other countries which have obtained substantially lower infant mortality rates. Over 98 per cent of births in the United States occur in hospitals and are attended by and large by physicians. In the Netherlands, where the infant mortality rate is 40 per cent lower than in the United States, the majority of births occur in the home and are attended by midwives at presumably considerably less cost. Other countries have also attained much lower infant mortality rates without the utilization of large scale hospital confinement for births or highly specialized physician manpower in attendance at most births.

Table 1

Age Specific Mortality: 1970

Age	Rate per 1,000
Fetal (1968)	15.8*
Neonatal	14.9**
Infant	19.8**
1-4 Years	.8
5-14 Years	.4
15-24 Years	1.3
56-64 Years	16.6

*Fetal deaths per thousand pregnancies

**Infant deaths per thousand live births

As Table 1 indicates, infant mortality rates are almost a natural choice as an index of health because they are the highest and therefore presumably the most stable age specific mortality rate for any age group in the United States younger than 65 years. In fact, the death rate during the first year of life is about equal to that for ages 1-29 combined. This high visibility has been another reason why critics have latched onto the relatively poor infant mortality showing in the United States as an indictment of the health care system despite the fact that infant mortality rates in this country are low - very low relative to less developed countries.

Defenders of the United States' infant mortality experience position relative to other countries originally marshalled two arguments to account for the relatively poor showing in the United States. First, it was maintained that statistics were kept differently in the United States than in other countries and that because of the much higher incidence of hospital/physician attended births in this country, births that resulted in deaths were much more likely to be reported as such (Chase, 1967). It was argued that this difference accounted for a substantial amount of the differences observed in the rates. A closer examination of the statistics indicate that while this is a valid explanation for part of the gap, it does not seem to adequately account for the trend in United States of a relative deterioration of infant mortality rates which has gone contrary to any indication of relative improvement in health care in this country.

A second argument advanced is that the United States has an extremely heterogeneous population and that therefore it is invalid to compare rates in this country with rates in small generally homogeneous European countries, presumably because there are genetic factors involved in the relatively favorable position of these other countries. Unfortunately this argument, whether or not valid, is once again an argument that would support differences in rates at an absolute level but is not conducive to supporting trends which have shown until recent years a definite deterioration in the position of the U.S. relative to other countries.

Historically, the relative deterioration of the U.S. position occurred during the decade from approximately 1955 to 1965 when the rate of decline in the U.S. infant mortality rate slowed to less than 10% for the decade while rates in other countries were dropping relatively rapidly. Since the period of the late 1960's, rates in the United States have begun to decline rapidly at a rate of decline approximating that experienced in other advanced countries in the earlier period. It is still too early to determine for the years much beyond 1970 whether or not the recent substantial declines in infant mortality in this country have reversed the trend of continuing deterioration in the U.S.'s relative position.

The relative dynamics of both cross sectional and secular trends in infant mortality clearly point out that rates are the result of a complex reaction of medical care, social, biological and behavioral factors and that any one explanation or approach to the subject is likely to fail because of its presumed inability to cope with the other factors involved. Despite the problem of determining causality, there has been a definite policy thrust to attempt to reduce infant mortality in the United States by increasing the amount of medical care rendered to the pregnant population. The rationale behind this approach has been strengthened by a study from the Institute of Medicine (1973), Infant Death: Analysis by Maternal Risk and Health Care, which seemed to imply fairly strongly that medical care had an important role to play in determining infant mortality. The most significant finding to emerge from the study was the projection that if all pregnant women in New York City in 1968 had received what was defined by the investigators as "adequate" prenatal care, the infant mortality rate could have been reduced by as much as a third. This lower rate would have represented a very respectable showing relative to contemporarily reported international rates. Moreover, the study found that there appeared to be a substantial misallocation of care resources amongst pregnant women when the relative risks associated with individual pregnancies were taken into consideration. Women who were defined to be at high-risk were less likely to receive any care while women who were defined to be at low-risk were more likely to receive "adequate" care.

Before plunging ahead with a policy designed to increase the utilization of medical services in order to reduce infant mortality rates, several subsidiary topics should be considered. One is the question as to what kind of information can we obtain about the selection process which leads to the finding that those women who presumably would have benefited most from care were least likely to receive it. Clearly this can be an indication of malfunction of the health care system but it may also be an indication of the absence on the part of the individuals of demand for care. This problem is in a sense highlighted by the study's own finding that in assigning women to high-risk groups on the basis of either "medical" or "socio-economic" risk that prenatal medical care had its most substantial impact in reducing infant mortality for those women who were judged to be high socio-economic risks but much less of an impact on those women who were deemed to be at high medical risk. The question naturally arises as to the rationale behind categorizing women as being at high socio-economic risk and whether there was an element of circular reasoning involved, i.e. groups of women who traditionally did not seek care had a previously documented history of bad outcomes and were therefore defined as being at high socio-economic risk. Secondarily, the question as to why these women did not seek care or obtain care and the implications of their failure for programs designed to extend care to them become very significant.

The whole area of household decision making has been another area that has opened up to study by economists, particularly within the last decade. Within this area, Grossman (1972, 1974) in particular, has focused on a household production model of the demand for health and subsequently the demand for child health and has attempted to determine those variables that are responsible for determining the demand for medical care and presumably the production of health. In addition, a number of economists have turned their attention to examining the determinants of family size and have viewed child services as being a significant commodity demanded and produced within the household. They have examined desired family size, desired child spacing and the demand for contraception and child attributes within the context of the household production model.*

It seems therefore that a fruitful extension of research on the infant mortality problem, with particular attention to the relative efficacy and utilization of prenatal services, would be to attempt to develop a model of demand for infant health and hence survival which utilizes the household production perspective. It seems logical to presume that the demand for healthy or surviving children is clearly derived from the demand for child services themselves and moreover, that the demand for medical care in the production of favorable pregnancy outcomes is derived from the demand for surviving children and the presumed efficacy of care in producing survival. This will be the approach that I take in this paper. I shall assume that there is a household demand for children which can be satisfied by successful pregnancies and that, in making a determination of the desired number of pregnancies and the expenditure per pregnancy, the household shall take into consideration

*See for example, Willis (1972), De Tray (1972), Michael (1972), Michael and Willis (1973), Ross (1974), Ben-Porath and Welch (1972), and Becker and Lewis (1972), to mention only a representative sample.

the cost of producing survivors and the post-pregnancy cost of children relative to the cost of other goods. In determining the demand for care and in attempting to measure the efficacy of medical care in the production of healthy infants, I shall view such demand as being derived from the demand for surviving children and concentrate on the importance of household experience with previous pregnancies in determining both the amount of resources to be expended on a specific pregnancy and the amount of care demanded during a specific pregnancy and attempt to accurately measure the outcome of such resource allocations in order to more clearly define the production process.

In the next chapter, I examine possible fertility responses on the part of households to completely exogenously determined mortality. The implications of various forms of household response for measured mortality rates demonstrate that measured infant mortality rates are a function not only of biologically and/or medically determined survivorship but also of the demand for additional pregnancies in the face of losses. In Chapter 3, I develop a model of decision making for expenditures on pregnancies based on prior experience with pregnancy. I attempt to integrate demand for children with the demand for pregnancies and the demand for expenditures on individual pregnancy based on prior experience and other exogenous variables.

Chapter 4 contains a description of an empirical model that can be estimated to test some of the implications of the theoretical model. The particular focus is on the pregnancy as a production process and the emphasis is on formulating a system of equations to describe sequential behavior in this context. Two demand relationships are formulated.

They measure demand for prenatal care by using both the interval to the first prenatal visit and the number of visits as dependent variables and two outcome measures are estimated as the result of the production process. They are birthweight and survival. /

Before empirical estimates of the relationships are presented, specific problems encountered in attempting these estimates and the solutions employed are discussed in Chapter 5. In Chapter 6, estimates of the demand functions for prenatal care are presented and discussed using cross sectional data from the 1970 New York City birth cohort. Using the same data set, outcome relationships are estimated using OLS regression techniques and discussed in Chapter 7. Since OLS may not be the appropriate technique to employ with dichotomous dependent variables, such as death or survival, full information maximum likelihood estimates of the logit of the death function are presented in Chapter 5 and compared with the OLS estimates. A summary of the conclusions drawn from the estimated relationships are discussed in Chapter 9.

Chapter II - The Relationship Between Replacement Fertility and Measured Infant Loss Rates

The notion that infant and child mortality rates will have a strong influence on family fertility decisions has played an important role in the prediction of population growth patterns by demographers. Observation of family responses to child losses and of high fertility rates in areas with high child mortality experience has led many to conclude that "replacement" is an appropriate model of behavior under these circumstances. By "replacement" is meant the almost automatic attempt to compensate for a dead offspring with a new birth. The basic assumption underlying such hypothesized behavior is that parents have a certain targeted family size and that they will strive to attain this goal even in the face of substantial losses. It is frequently assumed in such models that children represent a form of old age security for parents and are an important form of investment for this purpose and that a specific number are required to assure fulfillment of this function. If death rates are high throughout the entire childhood period, and even during young adulthood, then parents will require a large number of births to assure an adequate number of survivors. Concern has been expressed about the possibility of explosive population growth if declines in child mortality, which are likely to accompany economic development, are not matched by compensatory declines in the birth rate and improvement in conditions of child health lead parents to over shoot their goals.

Several economists working in the area of fertility determination have pointed out that this simplistic model of replacement does not adequately deal with the possible relationships between fertility and child mortality. DeTray (1972) has pointed out that since pregnancy is not costless, high infant loss rates may in fact discourage pregnancies as at the margin the family finds that the utility associated with the probability of the child's surviving is less than the disutility associated with maternal mortality (the only cost he explicitly cites as being associated with pregnancy) which is likely to be high in areas where child mortality is also high.

O'Hara (1972) has investigated the cost and return streams associated with investment in children at different stages of their life cycle. He uses a model in which family decisions about children involve the risk of death at each age rather than the overall survival rate. Using this technique, benefit and cost streams for children can be derived which reflect the expected survival experience of the individual family. An important aspect of this analysis is that it treats expenditures on children as endogenous and allows parents to adjust expenditures to reflect experience so that changes in the survival rate need not only influence the number of children desired but their average cost ("quality"). If parents are more likely to experience net benefits from children as the children survive longer, O'Hara points out that mortality declines, particularly after infancy, are likely to increase the expected net return from pregnancy and encourage parents to devote more resources to children. This process should also lead parents to substitute "higher quality" children for "lower quality" children. While the net result of the various substitution effects on both the birth rate and more importantly the rate of population growth are uncertain in this model and can only be resolved empirically, O'Hara's model does add a dynamic quality to family decision processes. Moreover, the model may have significant explanatory value in advanced societies where childhood mortality drops off sharply after infancy so that families have greater flexibility in planning for additional births. Household expenditures on child quality may then reflect the survival experience of individual offspring.*

Before developing my own approach to the problems of the demand for children and child survival, I would like to point out that a limitation of the works cited is that they are primarily concerned with explaining fertility behavior resulting from the health (survival) status of children and not with the determinants of infant health per se. Thus, the observed correlation between high child death rates and high birth rates in LDC's might alternatively be explained by the hypothesis that a high birth rate in the presence of extremely limited family resources results in high child loss rates as resources are spread too thinly. If family planning is not widely practiced in these societies, it would appear that the assumption of exogenous death rates and endogenous birth rates, which underly the simple replacement hypothesis, are not tenable. Moreover, neither DeTray nor O'Hara have considered the implications of

---

*An important distinction can be made between so-called developed and less developed nations based not only on the differences in the absolute level of mortality at all ages but in relative age specific mortality rates. As noted previously (Chapter 1, Table 1), in the more advanced nations, mortality rates for children over a year old are extremely low both absolutely and relative to infant mortality rates. In such circumstances, families will be able to attain desired family size by replacing only those desired children who actually die, since the death is likely to occur while the mother is still fecund. In less developed countries, where death rates are high throughout childhood, families have much less control over completed family size. A large number of births may be desired in such situations regardless of individual experience with survival during infancy because the risk of the loss of a child at five or ten years of age may still be substantial and unpredictable. The danger is that such a fertility strategy may lead to explosive population growth if parents fail to reduce the targeted number of births in response to improved exogenous conditions which might significantly reduce deaths among older children.

child survival being dependent on the amount of expenditures on children - although both have made expenditures conditional on survival. Lastly, little consideration has been given as to how the family forms its expectations as to child survival - while the assumption seems to be that societal values are applied in individual decision making, it seems likely that individuals have other sources of information about their own probable child rearing potential and that these may play an important role in their individual fertility related behavior. I shall defer to subsequent sections of this study the theoretical implications and empirical investigation of considering child survival as being endogenously determined. I shall first investigate some of the implications for aggregate infant loss rates of individual decisions regarding child bearing under conditions where pregnancies are costly, outcomes uncertain and loss rates are exogenous although not uniform throughout the population.

The model that I shall examine is a variation of one utilized by Ben-Porath and Welch to examine the effect of the unpredictability of certain child characteristics on fertility.* Thus they point out that, while many economists, following Becker (1960), have chosen to consider child "quality" as being a function of parental expenditures on children, there are attributes of children such as sex, genetically determined

---

*Ben-Porath and Welch have jointly and individually published four papers on this topic: Ben-Porath and Welch (1976); Welch (1974); Ben-Porath and Welch (1973); and Ben-Porath (1973). Each of their papers draws heavily on their preceeding work as does much of the theoretical discussion in this study. Several of the mathematical results derived herein are similar to those presented in Ben-Porath (1973), although they were developed independently as an extension of the Ben-Porath and Welch (1973) study which was my first contact with their work.

physican characteristics or intelligence and even survival which are beyond parental control but which will effect the amount of satisfaction parents obtain from individual children. While Ben-Porath and Welch have been concerned largely with the effect of the sex of children on parental child bearing decisions, I shall demonstrate how a similar model may be useful in explaining differences in the measured infant mortality experiences of different populations.

Consider a situation where families have consumption choices covering two goods and attempt to maximize utility subject to the usual income constraint. Thus,

$$U = U(C, S) \qquad (1)$$

where U is lifetime utility, C is surviving children and S is all other goods. Further, let

$$C = p \cdot n \qquad (2)$$

where n is the number of pregnancies and p is the probability of survival of any individual pregnancy (which we shall call reproductive efficiency). We may then define the lifetime income constraint as

$$Y = S + n \cdot \pi_b + pn \cdot \pi_c \qquad (3)$$

where all values have been fully discounted and S = total expenditures on other goods; n and p are as defined in (2); π_b is the price of an individual pregnancy including both the direct costs such as medical costs and the indirect costs such as the opportunity cost of the mother's time lost during pregnancy and perhaps the inputed cost associated with increased risk to her health resulting from pregnancy; and π_c is the cost of a surviving child including both direct costs and parental time costs involved in raising a child. If we substitute the expected values of C from (2) and Y from (3) into the utility function and call the new utility function of the expected outcomes U_E, we may maximize U_E with respect to the number of pregnancies subject to the income constraint and get

$$\frac{\partial U_E}{\partial n} = p(U_c) - (\pi_b + p\pi_c)\, U_s = 0 \qquad (4)$$

where $U_c = \frac{\partial n}{\partial c}$ and $U_s = \frac{\partial U}{\partial s}$. Not surprisingly, this indicates that women should continue to become pregnant until the expected utility of an additional surviving child is equal to the expected utility foregone due to the decreased consumption of other goods.*

---

*Note that maximizing U_E is not strictly equivalent to the traditional concept of maximizing expected utility $E(U(C,S))$. However, p is exogenous and random so that one might want to pursue an expected utility analysis by incorporating the density function of the survival parameter p, $f(p)dp$. Such an approach is presented by Ben-Porath (1973) and adapting his formulation to the model considered above yields,

$$E(U(C,S)) = \int f(p)U(pn, y-\pi_b n - pn\pi_c)\,dp. \qquad (4a)$$

While (4a) can be maximized with respect to n, the interpretation of the result depends on consideration of parental attitudes toward risk as captured in the specific form of the utility function. It does not appear that pursuing the formulation expressed in (4a) will yield any additional strong insights without substantially complicating the analysis and calling for additional assumptions. Moreover, we can demonstrate that the family formation rule expressed in (4) is actually not as restrictive as it may appear from its development above.

The relationship expressed in (4) can be derived within an expected utility framework if we modify the ground rules slightly. Thus, let us consider the utility associated with an additional pregnancy to a family that already has experienced n_o pregnancies and has C_o surviving children. If p is the probability of survival and π_b, π_c, and S are as defined previously, and if there are two possible outcomes of the $(n_o+1)^{st}$ pregnancy, then the expected utility associated with the $(n_o+1)^{st}$ pregnancy may be written as

$$E(U_{(n_o+1)}) = E(U_n) = p(U(C_o+1, S+(n_o+1)\pi_b+(C_o+1)\pi_c) + (1-p)(U(C, S+(n+1)\pi_b+c\pi_b). \qquad (4b)$$

If we allow U_c $(=\frac{\partial U}{\partial C})$ and U_s $(=\frac{\partial U}{\partial S})$ to denote small changes in U associated with changes in C and S around the initial endowment point $(C_o, S+n_o\pi_b+C_o\pi_c)$, then

$$E(U_n) = p(U_c - U_s(\pi_b+\pi_c) + (1-p)(-U_s\ _b) = p(U_c) - U_s(\pi_b + p\pi_c) \quad (4c)$$

which is exactly the same as (4). Moreover, (4c) may be interpreted as a decision rule in a manner analogous to 4, that is, so long as $E(U_n) > 0$

Before examining the implications of this model for infant death rates I should like to make explicit some implications of the assumption that p, π_c and π_p are all exogenous. Although one might consider models where families trade-off π_b for π_c, or, as I shall consider subsequently, the possible trade-offs between p and π_b, here I regard them all as given for a particular family. An advantage of this approach, aside from simplicity, is that it results in an unambiguous definition of p, reproductive efficiency, as the probability that an individual pregnancy will survive infancy for a given level of π_b. Thus changes in p that I shall consider below will be independent of the level of π_b and may, for example, represent changes over time or differences in a cross section resulting from genetic differences, differences in maternal health, or non-pregnancy specific environmental differences.

In a world of certainty, p may be treated similarly to a technological change parameter and the response of n to changes in p derived in a straightforward manner. Thus, let

E = the percent change operator, and

$e_{a,b}$ = the elasticity of a with respect to b

then from (2)

$$En = EC - Ep = e_{c,k}\, Ek - Ep \qquad (5)$$

where $k = \frac{\pi_b}{p} + \pi_c$, the cost of a surviving child. Now,

---

(footnote continued)

it is desirable to have an additional pregnancy, but if $E(U_n) < 0$, it is no longer desirable to have another pregnancy (clearly, one is indifferent at $E(U_n) = 0$). Note that as one moves along the budget constraint, increasing n, for fixed p, π_b, π_c, and Y, U_c should fall relative to U_s so that $E(U_n) = 0$ should define the point of maximum utility as in (4).

$$e_{k,p} = -\frac{\pi_b}{\pi_b + p\pi_c} = -\alpha \tag{6}$$

where α = the share of pregnancy cost (π_b) in the total expected cost of a pregnancy ($\pi_b + p\pi_c$). Then, substituting (6) into (5) and gathering like terms, yields

$$En = e_{c,k}\, e_{k,p}\, Ep - Ep = -(\alpha e_{c,k} + 1)\, Ep \tag{7}$$

$$\text{or} \quad e_{n,p} = -(\alpha e_{c,k} + 1) \tag{8}$$

$$\text{or} \quad e_{n,p} = -(\alpha e_{c,k} + 1) \quad (8)$$

If children are normal goods $e_{c,k}$ should always be negative while α will be positive, so that the relative size and even sign of $e_{n,p}$ will depend primarily on the price elasticity of children, $e_{c,k}$.*

We can examine the implications for different values of $e_{c,k}$ on $e_{n,p}$. Note, that if $e_{c,k} = 0$, then $e_{n,p}$ will equal 1 and we will be looking at the world of complete replacement. If $e_{c,k} < 1/\alpha$, the most likely situation, then a rise (fall) in p will only partially decrease (increase) n. Only if $e_{c,k} > 1/\alpha$ will an increase in p increase c and leave n unchanged. The rational behind these relationships rests on the dual function that p is made to play in this model. Thus, a rise in p causes the price of a surviving child (the desired commodity) to decline so that more children should be desired. However, an increase in p also means that fewer births are required per survivor so that the net effect on the number of desired pregnancies depends on the relative size of the parameters as indicated by the inequalities above.

Welch and Ben Porath (1973, 1976) are primarily concerned with the effect of differences in the sex composition of children on completed family size. They examine situations where there is uncertainty as to the true value of λ, the proportion of boys and are particularly concerned about the implications of the different ways in which expectations

---

*$e_{n,p}$ will, of course also depend on the relative magnitude of α however, it would appear that α is likely to be quite small and estimates of its size are not available while estimates of some elements of $e_{c,k}$ have been attempted (Willis, 1973).

about λ are formed.* They consider cases where λ is based on only prior information (the dogmatic case), only on the experience of the individual family (the naive case) and on some weighted average of prior information and personal experience.

In the dogmatic case, couples whose actual experience deviates from their expectations will tend to have larger families as it is not unlikely that they will view the loss associated with being "unlucky" as an income loss rather than as an indication of a difference in the potential sex mix of their children. In the case where families are "naive" and expect that the sex composition they have already experienced will repeat itself, they are less likely to have larger families even if they have experienced an undesirable sex ratio because they will view the unfortunate experience as an indication that they are faced with a higher price for future children of the desired gender. Needless to say, in the intermediate case, "unlucky" couples will respond somewhere between the two extremes but will still tend to have larger families. In all situations, the critical empirical decision parameter is the price elasticity of demand for "quality corrected" children.**

---

*λ in their formulation is equivalent to p in our discussion of mortality for most of the formal implications of the model. The primary difference would appear to result from the fact that a surviving child of the less desired sex would probably be more costly than an infant lost at birth - however, the survivor may yield some utility so that the net cost would be difficult to calculate unambiguously.

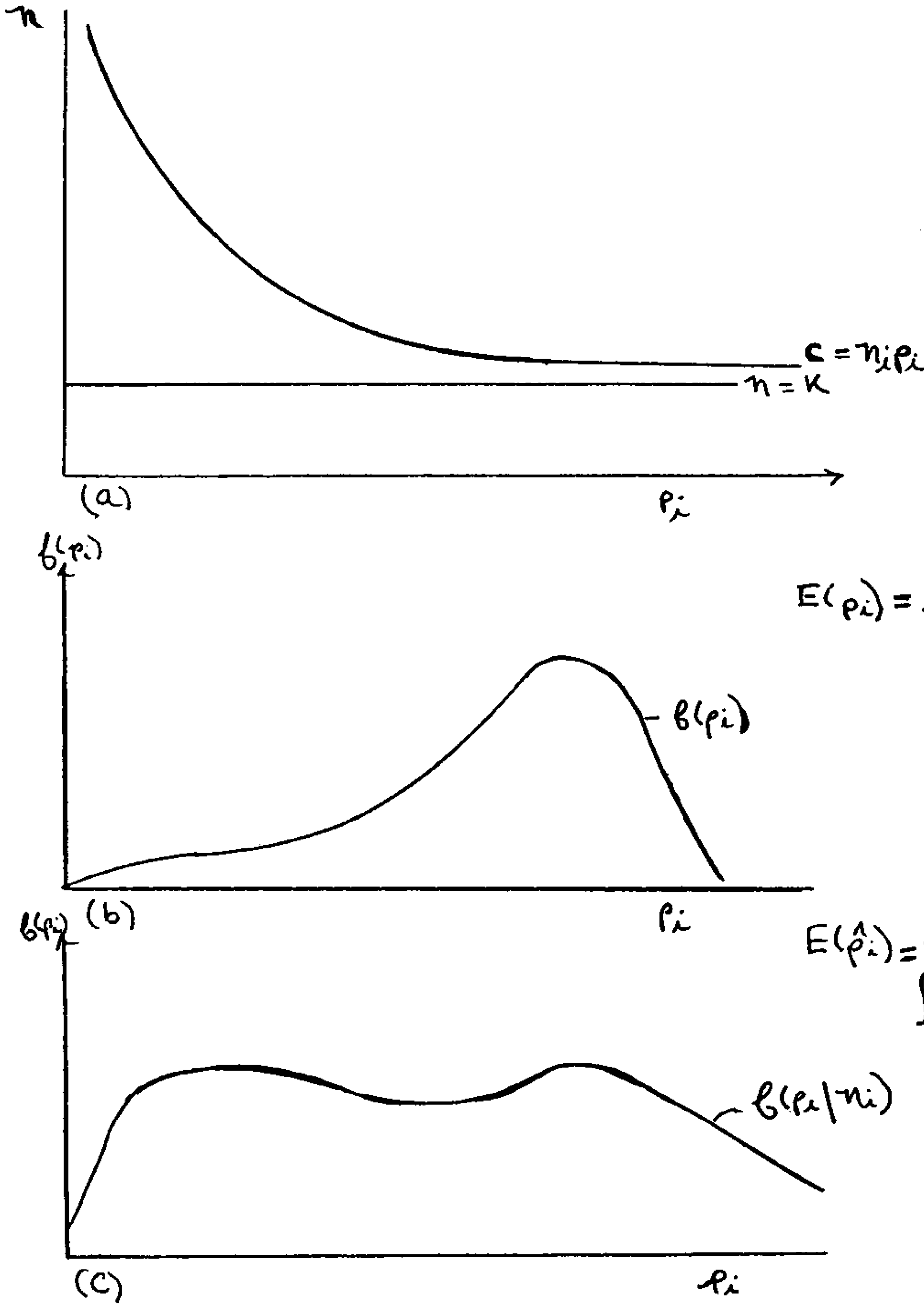
**Ben Porath (1973) has subsequently pointed out, that extension of the model to allow for "pure risk" considerations, as p or λ is not only exogenous but also random, does not yield preductions which are independent of the type of utility function assumed and that aside from possible risk reducing portfolio effects which are fertility inducing, few other straightforward testable hypotheses can be inferred despite considerable complication of the model.

In the model I have presented above where the important element of uncertainty is survival and the cost of pregnancy is only a small part of the total cost of having children, it is quite likely that the tendency will be for "unlucky" couples (those with high infant loss experience) to demand more pregnancies although their completed family size may be no greater (and probably smaller) than "lucky" couples.

Let us now consider the probable effects of these attempts at replacement on measured infant mortality rates in different populations. The implications of such behavior for mortality rates are best seen graphically in Figure 1. Here, I have contrasted the measurable results of two somewhat extreme decision rules parents may follow with regard to pregnancies and infant losses. Following the first rule, parents decide on a certain number of pregnancies, K, regardless of the outcome of the pregnancies; such behavior is illustrated by the straight horizontal line $n=K$ in panel (a). On the other hand, if parents set a target family size of C and p_i (reproductive efficiency) is distributed randomly throughout the population then the relationship between n_i (the number of pregnancies per family) and p_i (individual survival experience) will be represented by the hyperbola $C = n_i p_i$, as also shown in panel (a).

If p is distributed in the population as indicated in panel (b), then under the first decision rule, $n_i = K$, the measured mortality rate for the population will equal the mean of the underlying distribution for that population. Let us call $f(p_i)$ the frequency of any given p_i in the population. Then for each level of p, there will be $f(p_i)\cdot K$ births and $p_i f(p_i)K$ survivors. Thus, we can calculate the survival rate:

Figure 1 (II). The Effect of Replacement on the Distribution of Births by Reproductive Efficiency and the Survival Rate.



$$E(\rho_i) = \int_0^1 \rho_i\, b_{\rho_i} d\ldots$$

$$E(\hat{\rho}_i) = \frac{1}{\int_0^1 \frac{b_{\rho_i}}{\rho_i} d\rho_i}$$

$$\text{Survival rate} = \frac{\text{Survivors}}{\text{Total births}} = \frac{\int_0^1 Kp_i f(p_i)\, dp_i}{\int_0^1 Kf(p_i)\, dp_i}$$

$$\int_0^1 p_i f(p_i)\, dp_i = E(p_i). \tag{9}$$

If families follow the second rule and replace all losses entirely then $C=p_i n_i$ and the distribution of births by reproductive efficiency will be as pictured in panel (c). Here for each $p_i, n_i = \frac{C}{p_i}$ and hence, total pregnancies in the population will equal

$$\int_0^1 \frac{C}{p_i} f(p_i)\, dp_i \tag{10}$$

and total survivors

$$\int_0^1 p_i \left(\frac{C}{p_i} f(p_i)\right) dp_i \tag{11}$$

The survival rate will be

$$\frac{\text{Survivors}}{\text{Total Births}} = \frac{1}{\int_0^1 \frac{f(p_i)}{p_i}\, dp_i} \tag{12}$$

Let us call this rate, $E(p_i | C)$.

Infant loss rates will be higher under the hypothesized replacement rule than the equal number of pregnancies rule. This can be inferred from a straightforward examination of the values of integrand within each integration. For each value of p_i, the integral of $E(p_i)$ has the value $p_i f(p_i)$ which is less than the value associated with that same p_i under $E(p_i | C)$, $\frac{f(p)_i}{p_i}$, since $0 < p_i < 1$. Hence, over all values of p

$$\int_0^1 \frac{f(p_i)}{p_i}\, dp_i > \int_0^1 p_i f(p_i) dp_i > \frac{1}{\int_0^1 \frac{f(p_i)}{p_i}\, dp_i} \tag{13}$$

and $[1 - E(p_i)]$ the infant loss rate under rule one is less than

$[1 - E(p_i|C)]$ the infant loss rate under the second rule (complete replacement)

Given that the infant loss rate will be higher under the decision rule requiring complete replacement, it is imperative that we investigate how likely this decision rule is. Recalling the previous discussion of the relationship between $\ell_{c,k}$ and α, note that if p is high on average and π_c substantially greater than π_b then α will be small and $\ell_{c,k}$ will have to be quite high to discourage substantial replacement. In the possible learning examples discussed by Welch and Ben-Porath, only the extreme situation of "naive" learning (where people assume that their past experience is the only predictor of future results) will the tendency to replace by "unlucky" (those with biologically low p's) be restrained. In the "dogmatic" case, they are not influenced by experience and in the situation of moderate learning, they may require many pregnancies before deciding that they are indeed biologically "unlucky."*

Some evidence that people tend to replace and that p may actually be distributed randomly in the present U.S. population can be found in Table 1.** This table reports the previous pregnancy experience of all

---

*This is the most likely situation for families in the extreme lower end of the distribution of p. Clearly, they cannot physically (as well perhaps as economically) sustain the numbers of pregnancies required by the decision rule $c = n_i p_i$ as $p_i \to 0$. These women may follow a maximum n_i rule or stop becoming pregnant after the first successful birth (see Billewicz [1973], for evidence of this behavior).

**It is important to distinguish between observed behavior and the true variation in p in a cross section. If p did not vary and infant mortality was determined by a simple Bernoulli process, then even in the absence of behavioral response to loss (e.g. under the rule C = K) the distribution of observed p would follow a binomial distribution. Under such circumstances, individual learning from past experience would have little value to the family, although individuals might indeed act as though experience had informational value. Notice, moreover, that even if families follow the rule requiring complete replacement, the aggregate measured mortality rate for a given time period will not be affected since their probability of success on subsequent pregnancies is independent of their past history.

Table 1

Cumulative Experienced Pregnancy Loss Rate of Women with at Least One Prior Pregnancy in the 1970 New York City Birth Cohort by Outcome of Indexed Pregnancy

Outcome of 1970 Pregnancy	Experienced Child Loss Rate Per 1,000 Births	Experienced Fetal Death Rate Per 1,000 Pregnancies
Infant Death	57.5	161.9
Late Fetal Death (Stillbirth)	43.0	197.0
Surviving Infant	32.0	123.0
1970 Rate based on all pregnancies	21.8	*
Average of Rates 1955-1969	25.5	132.1

*1970 Fetal Death Rate not calculated because of very poor quality of recording of this data on Fetal Death certificate due to legalization of abortion in New York City in July 1970 (see page ___).

members of the 1970 New York City live birth and late fetal death cohort. Two facts stand out in this tabulation. First, the death rate experience of members of the 1970 cohort is higher than the average experience of New York City birth cohorts over the previous decade - this would tend to indicate a form of replacement behavior since it might be inferred that other things equal, members of the previous cohorts who have been "luckier" have dropped out having reached their desired family size earlier. Second, there is a strong propensity for infant survival experience to repeat itself within the population - thus those women who experienced an infant death in 1970 have a significantly higher historical loss rate than those who had a surviving infant in 1970. Moreover, this repetition of past behavior is even selective as to the type of pregnancy loss - e.g., the historical fetal death rate is higher for women who have experienced a fetal death in 1970 than for those who experienced an infant death, while infant death rates are higher for those in the 1970 infant death category; rates for both types of losses are higher for those "unlucky" in 1970 than for those who bore surviving children in 1970.

The implications of the model advanced above as well as the evidence presented in Table 1 would seem to suggest that there is a tendency for individuals even in modern societies to replace lost pregnancies and moreover, that there appears to be a tendency for experience to repeat itself (the culmulative loss rate of those in the 1970 infant death group approaches a 15% infant mortality rate compared with a 2% rate overall in 1970). In light of such behavior, it would appear that differences in the exact distribution of p, measured perhaps by the higher moments of the distribution, for individual population groups may significantly affect measured infant loss rates in a manner not adequately captured by variations in $E(p_i)$ alone. Unfortunately, it is not possible to predict, in general, the expected effect of differences in the distribution of p on $E(p|c)$ (our approximation to the measured mortality rates given replacement) by reliance on the moments of the distribution. The mathematical stumbling block is that it is not possible to infer anything generally about movements in $\sum_i(\frac{1}{X_i})$ (the general form of the denominator of $E(p|c)$) from information about the movement of $\sum_i X_i$ (or $\sum_i X_i^2$, etc., the general form of the moments of $f(p_i)$).* One is, therefore, left with the relatively weak statement that measured infant loss rates should depend on the entire distribution of p_i not only its mean but I am unable to predict the expected direction of this relationship.

Some Empirical Evidence

Despite the fact that I am generally unable to predict the direction or magnitude of the expected effect of variation in the moments of the distrib of p on measured infant surrival or mortality rates, it would appear worth examining whether such effects can be detected empirically. The value of such an endeavor stems partly from the reasonableness of the assumption of at least partial replacement behavior by most prospective parents and partly from the importance of measuring the extent of the effect of behavior not directly related to the health care system on a widely utilized health index such as infant mortality. If cross sectional and secular variations in measured infant mortality rates are due to differences in the underlying distribution of p as well as differences in the demand for surviving offspring (as captured in differences in $e_{c,k}$), greater care must be

---

*This topic is explored more fully in the Appendix to this chapter.

exercised in using measured infant mortality rates as a general health indicator and different policy alternatives may provide otherwise unexpected results in affecting the measured level of infant mortality in a given area or for a particular population group.

As an example of how consideration of the fertility/mortality interaction model might help in interpreting changes in measured infant mortality rates, consider the relatively rapid decline in infant mortality rates in this country since the late 1960's. In 1950, the U.S. infant mortality rate stood at 29.2 deaths/1,000 live births and the U.S. ranked sixth lowest among 15 industrialized nations.* By 1955, the U.S. rate had declined to 26.4 per 1,000; however, by 1965, the U.S. rate had only declined to 24.7 per 1,000 live births - a decline of only 6% in ten years. In contrast, the Japanese rate dropped from 39.8 to 18.5 over the same decade, a decline of 54% and even the Swedish rate, already the lowest among all nations in the world, declined 24% during the same period. By 1969, the U.S. ranked 15th among the same 15 industrialized nations by the measured infant mortality rate and critics of the U.S. health care system pointed to the relatively high U.S. rate, and slow rate of decline as evidence of a malfunction in the U.S. health care system. In the late 1960's the U.S. rate began to fall and it had reached 16.7 per live births in 1974 (Wegman, 1975), while still high by international standards, champions of the U.S. health care system could point to the 32% drop in nine years as evidence that the system was functioning well. While it would be too simplistic to suggest that the simple model elaborated above in which measured mortality rates depend partially on fertility

*All pre-1970 rates and rankings from Chase (1972).

behavior can provide a complete explanation for this reversal in the U.S. trend, the model does suggest that failure to consider the fertility/mortality trade off can lead to misinterpretation of such trends and particularly to an overly "health system" oriented response to changes in these measured rates.

For one thing, the U.S. has also experienced a rapid and persistent decline in the birth rate which also began in the late 1960's - roughly coincident with the decline in the infant mortality rate experienced since then. While a check of the simple replacement model indicates that measured mortality rates ($E(p|C)$) should be independent of desired family size (C), such a change in the fertility rate may indicate a change in basic fertility behavior and in particular a change in $\epsilon_{c,k}$ which will effect measured infant death rates under the replacement hypothesis. Innovations which reduce the cost of family planning, and particularly pregnancy prevention, such as the oral contraceptives introduced in the 1960's and the legalization of abortion in the 1970's, will tend to increase $\epsilon_{c,k}$, the elasticity of demand for children with respect to their price.* Such innovations should also increase the income elasticity of the demand for children as in both cases technological change reduces the cost of making a change in desired C (or n for that matter) and therefore should make changes in C more responsive to changes in other variables, such as income, in the families' decision matrix. Referring to equation (3), we note that an increase in $\epsilon_{c,k}$ (so long as $\epsilon_{c,k} < 1/\alpha$) should lead to a reduction in $\epsilon_{n,p}$ and an overall reduction in th[e] tendency to replace unsuccessful pregnancies. This reduction in the tendency [to]

---

*Where the cost of preventing a birth is not, properly, considered a cost of having a child.

This hypothesis was tested using 1960 cross-sectional data for the U.S. where the units of observation (specific population groups) were SMSA's. The dependent variables were neo-natal and post-neonatal infant mortality rates, for whites and blacks in the 59 largest SMSA's for which it was possible to distinguish the independent variable, a surrogate for reproductive efficiency, for whites and non-whites and a substantial number of non-whites were blacks. The infant mortality rates were averaged over three years, 1959 to 1961 and centered on 1960. As the independent variable, the proxy that was chosen for reproductive efficiency was completed years of schooling (education) of females in the SMSA's aged 14-44.

There are several good reasons for choosing education as a proxy for reproductive efficiency as we have defined it. For one thing, several authors have recently stressed the importance of education in potentially increasing efficiency in household production.* In particular, Grossman (1972) has stressed the importance of education in raising the efficiency with which the health of an individual can be produced within the household. It is very likely that such an effect would carry over to the production of offspring. Moreover, the fact that during the pregnancy the mother's body is the means of production of the offspring and the mother's health an important factor in determining the health and survival of the child would indicate that a variable such as mother's education which is highly correlated with health, no matter what the direction of causality, could be a valid instrument for measuring reproductive efficiency.

Along these same lines, it might be argued that measured education is a fairly good index of the amount of investment that has been made since childhood in an individual and that such an investment may pay off in the

---

*See for example, Micheal (1972), Becker (1965, 1971), Gronau (1973, 1974) and Liebowitz (1974) to cite only a few.

replace, particularly by those with low levels of reproductive efficiency should lead to a reduction in measured infant mortality rates that is independent of any changes in the underlying distribution of p in the population, of any changes in the maternal/infant health care system or even of any changes in c, desired family size.*

A Specific Test

The potential importance in evaluating differences in measured infant mortality rates of fertility related reactions to different infant mortality experiences encouraged the further empirical investigation of the relationship implied in the replacement model. Specifically, I attempted to test the hypothesis that differences in measured infant mortality rates in a cross section of specific population groups would be a function not only of differences in mean reproductive efficiency among the groups, but also of differences in the distribution of p within each group as measured by the moments of the distribution.

---

*Similar conclusions have been advanced by Billewicz (1973) based on a study of the reproductive histories of 4,948 married women who had their first pregnancies in 1949-54 and who were followed up until 1964, i.e., for 10 to 15 years. He points out that there appears to be a pattern of "selection by success," i.e. those women with successful pregnancies reach completed family size earlier and drop out of the cohort (a pattern suggested by the data in Table 1), leaving a greater concentration of women with low reproductive efficiency in the birth cohort. He points out that more effective contraceptive methods may stimulate this process and that interpretation of vital statistics during a period of flux caused by the introduction of more effective contraceptive methods can only be done with great caution. As an example of the possible confusion resulting from this change in reproductive habits, he contrasts the decline in the aggregate perinatal mortality rate for Scotland between 1968 and 1970 (nearly half of the reduction is attributable to a shift in the parity distribution of births) with his estimates of perinatal mortality rates by parity based on his estimate of increased rates of "retirement". His estimated party specific rates show a large increase in measured perinatal mortality rates for women having their third and fourth pregnancies presumably because reduced costs of pregnancy prevention have increased the drop out rate of mothers with earlier successes, i.e., those with higher levels of reproductive efficiency. This would also tend to suggest that there is a pure birth order effect leading to a higher loss rate among previously successful reproducers as parity increases.

production of offspring, either because the mother is healthier herself or because she is better physically suited to carrying fetuses (see Stearns (1958) for an argument along these lines.)

A specific benefit which derives from using the distribution of education of all potentially fecund females as opposed to using a variable relating specifically to pregnant women is that I hypothesize different retirement rates based on experience when households tend to replace and therefore, would expect differences in the distribution of reproductive efficiency, as measured by the moments of distribution, to be reflected in differences in the measured loss rate. If the successful at the upper end of the p distribution retire from child bearing, while the less fortunate continue, the explanatory variable should measure the relative importance of these groups in the total underlying population. Therefore it is desireable to use the distribution of education for the entire population as opposed to the distribution of education for those who actually had children in those three years, 1959-1961.

The complete distribution of female education is available in the 1960 Census of Population published in separate volumes for each state (Table __ for SMSA's within the State) categorized by "no education" and by years of schooling completed, through seventeen or more. The data are also further broken down by age into four broad age groups: for our purposes, 14-24, 25-29, 30-34, 35-44; and further categorized by race, white or non-white. Using this data, the relationship between measured neonatal and postneonatal mortality rates and the first three moments of the distribution of potentially fecund females by level of education was estimated by using OLS regressions. The independent variables are defined as follows: ED is the mean level of female education in the SMSA, VAR is the variance of the distribution of education and SK is the measure of the skewedness of the distribution. Since it is not unlikely that the distribution of years of schooling completed will be related to the age of the population, all the moments have been standardized to the mean age distribution in all sampled SMSA's by the direct method of standardization.

The results of regressing neonatal and postneonatal mortality rates on the moments of the distribution of female education are presented in Table 2. Separate regressions were estimated for blacks and whites. Of significance in this table, primarily since we cannot a priori predict the expected signs of the coefficients (except perhaps for the mean), is that it appears that the higher moments have significant explanatory power in these regressions as presented. In particular, in the case of neonatal death rates which have traditionally relatively poorly explained by socio-economic variables, of which female education might be considered one, it appears that the addition of the variance and skewedness of female education for both whites and blacks to the regression equation increases the R^2 substantially - from .07 to .23 for whites and from .03 to .24 for blacks. In the regressions where post-neonatal death rates are the dependent variable, although the impact of skewedness does not seem to be significant, the variance of the distribution not only has a significant coefficient but also makes substantial contribution to the R^2 of the regression.

Of substantial interest is that despite the substantial differences in mean mortality rates for whites and blacks (see Table 3) the signs of the coefficients of the moments for whites and blacks are the same for each age specific mortality rate. That is, the signs of the coefficient of ED in all cases are positive, for neonatal death rates the signs of the coefficient of VR are negative, while for post-neonatal death rates the signs of the coefficients of VR for both whites and blacks are positive and for SK the signs of the coefficients for both race groups are positive

Table 2

Regressions of Infant Death Rates on Moments of the Distribution of Female Education, Ages 14-44, for 59 SMSA's (1960)
(t-statistics reported in parenthesis below coefficient estimates)

	Dependent Variable: Neonatal Death Rates[1]						Dependent Variable: Postneonatal Death Rate[1]					
	White			Black			White			Black		
Variable	a	b	c	a	b	c	a	b	c	a	b	c
ED	-.91 (-2.10)	-1.19 (-2.43)	-.61 (-1.24)	-1.12 (-1.25)	-3.08 (-3.01)	-4.85 (-3.44)	-.69 (-2.31)	-.33 (-.99)	-.40 (-1.11)	-3.32 (-6.38)	-2.09 (-3.69)	-1.94 (-2.36)
VR		-.14 (-1.19)	-.60 (-3.20)		-1.63 (-3.48)	-1.32 (-2.67)		.19 (2.34)	.25 (1.82)		1.03 (3.86)	1.00 (3.47)
SK			4.08 (3.04)			8.61 (-1.74)			-.51 (-.53)			.71 (.24)
Constant	27.23	31.36	25.58	39.29	71.93	83.37	12.83	7.38	8.11	44.64	24.10	23.16
R^2	0.07	0.10	0.23	0.03	0.20	0.24	0.09	0.07	0.17	0.42	0.54	0.54

[1]Per 1,000 live births

Table 3

Means and Standard Deviations
of Dependent and Independent Variables Included in the Regression Estimates

	White		Black	
Dependent Variables	Mean	Standard Deviation	Mean	Standard Deviation
Neonatal[1] Mortality Rate*	16.94	1.03	28.20	3.54
Postneonatal[2] Mortality Rate*	5.02	.72	11.84	2.65
Independent Variables				
ED - mean education (years of schooling)	11.28	.30	9.88	.51
VR - variance about the mean	6.88	1.22	8.16	1.09
SK - measure of skewedness	.60	.19	-.41	.16
MEDSQ - mean education squared	123.72	6.55	98.18	9.98

*Per 1,000 live births

[1]Deaths occured at ages 0-28 days

[2]Deaths occured at ages 29 days-1 year

and close to significance for the neonatal but negative and totally insignificant for post-neonatal deaths. Since there is substantial similarity in the causes of death of whites and blacks in these two age categories, the finding of some consistency across racial groups on the significance and sign reversals of the coefficients of the moments would tend to support the notion that mothers' education is a good surrogate for reproductive efficiency and that the hypothesized relationship may be stable for different genetic and socio-economic groups.

An alternative explanation as to why the moments worked relatively well (particularly the variance as indicated by its explanatory power) in these relationships is that in fact the linear relationships we have specified between the mean education level and the mortality rate is incorrect and that the true relationship is in fact approximated by a quadratic or higher order polynomial. Thus, if an incomplete quadratic has been fit to the data, this might substantially increase the explanatory power of the moments of the distribution.

Evidence of such a possible relationship between a quadratic function and the moments of the distribution can be seen in the following set of equations.

Let the infant mortality rate (IM) for an individual be a quadratic function of maternal education (ED_i), i.e.,

$$IM_i = a + b(ED_i) + c(ED_i)^2 \qquad (14)$$

then for an entire population, n,

$$IM = \sum_i \frac{IM_i}{n} = a + b \sum_i \frac{ED_i}{n} + c \sum_i \frac{(ED_i)^2}{n} \qquad (15)$$

Now, given the expression for the variance of ED_i

$$\sigma^2_{ED_i} = \sum_i \frac{(ED_i)^2}{n} - (\overline{ED})^2 \qquad (16)$$

where $\overline{ED}$ is mean maternal education, and substituting (16) in (15) yields

$$IM = a + b\overline{ED} + c\sigma^2 ED_i + c(\overline{ED})^2 \qquad (17)$$

which is similar to the expression we have fit to our data previously except that the term $c(\overline{ED})^2$ is added. This manipulation yields the rather strong testable hypothesis that if the equation (17) is estimated the estimated coefficients of $\sigma^2 ED$ and $(\overline{ED})^2$ should be identical. In order to investigate whether or not this alternative hypothesis is at least reasonable and may potentially negate our conclusion regarding the importance of the moments, regressions on the same sample were run including the mean female education squared as well as the moments as independent variables. These are presented in Table 4.

We notice in Table 4 that the sign of female education squared is the significant for white neonatal, white post-neonatal and black neonatal mortality and insignificant for black post-neonatal mortality. Moreover, the inclusion of education squared changed the sign of the coefficient on mean education for whites but not blacks. It would appear to have substantially increased the R^2 in all situations where it has a significant coefficient. However, our expectation based on (17) that the coefficients of education squared and variance should be the same or similar are not borne out by these regressions and in fact, for three sub-groups the coefficients are of opposite sign. While this does not provide adequate grounds for rejecting the hypothesis that the initial specification was totally inadequate, it would appear that at least the quadratic form as derived in (17) is ruled out as an alternative explanation and there is some reason to believe that the moments are of value in explaining infant mortality.

In conclusion, we note that if people do tend to replace as hypothesized in the model, behavior for which there is some empirical evidence, then measured mortality rates for infants will tend to be related not only to mean reproductive efficiency but to the entire distribution of

Table 4

Regressions of Infant Death Rates
on Moments of the Distributions of Female Education
and Mean Education Squared,
Ages 14-44, for 59 SMSA's (1960)

Variables	Dependent Variable: Neonatal Death Rate[1]		Dependent Variable: Postneonatal Death Rate[1]	
	White	Black	White	Black
ED	-37.22 (-2.94)**	36.65 (2.05)*	-40.55 (-4.96)	3.58 (.33)
VAR	-.53 (-2.95)	-1.08 (-2.23)	.33 (2.85)	1.03 (3.48)
SK	2.35 (1.68)	-6.64 (-1.37)	-2.41 (-2.68)	.97 (.33)
MEDSQ	1.67 (2.89)	-2.10 (-2.33)	1.84 (4.91)	-.28 (-.51)
Constant	232.00	-121.40	234.50	-4.11
R^2	.33	.31	.43	.54

*Significant at 5% level

**Significant at 1% level

[1]Per 1,000 live births

N.B. t-statistic reported in parenthesis below coefficient estimates reproductive efficiency and that differences in the propensity to replace as well as differences in the entire distribution of reproductive efficiency may potentially influence measured mortality rates even though mean reproductive efficiency doesn't change nor does the state of provision of medical care nor other factors such as public health programs that might be of importance in determining infant mortality.

In the next section, we will consider the effect of expanding the simple model we have presented in this section to allow expenditures on pregnancy to be endogenously determined and examine the potential effect of the experience with prior pregnancies on not only the decision to have a subsequent pregnancy, but on the decision as to how much to expend in attempting to insure the success of the subsequent pregnancy if undertaken.

Chapter III- A Model of Decision Making on Expenditures per Pregnancy

In the previous chapter, we considered a model of the demand for children when the outcome of a particular pregnancy is uncertain. We demonstrated that the demand for surviving children as developed in the previous model leads to a derived demand for pregnancies and that given the interaction between pregnancy losses and the desired stock of children that replacement type behaviour can cause a systematic empirical relationship to be observed between measured infant mortality rates and the underlying distribution of reproductive efficiency among different population groups. We touched briefly on the probable effects of a history of "good" or "bad" luck on individual fertility behaviour where we allowed the individual decisions as to whether to continue to have children to be the only choice parameter the family faced. In this section, we shall expand the model to allow expenditures on a given pregnancy to be an endogenous variable that the family may manipulate. We will examine the affects of income, education, experience and wantedness of the particular child on the expenditure of productive resources and the probable pregnancy outcome.

Once again we assume a two good utility space where for the individual family

$$U = u(C,S) \tag{1}$$

where C = surviving children and S = the composite all other goods with price equal to 1. As previously C is a function of n = the number of pregnancies, however we assume that p = the probability of survival is a function of given reproductive efficiency RE which varies in the population but is fixed for the individual (although not known with certainty) and EX = expenditures on a particular pregnancy for such items as prenatal care, hospital care, special food etc. Thus

$$C = P(RE, EX) \cdot n \tag{2}$$

where $\frac{\partial P}{\partial RE} > 0$, $\frac{\partial P}{\partial EX} > 0$, $\frac{\partial C}{\partial P} > 0$, $\frac{\partial C}{\partial n} > 0$.

In addition to expanding our production function for surviving children we shall also expand our cost of children specification. Since we are concerned with expenditures designed to insure the survival of the infant, we will assume that K, the cost of a surviving child is the same for all families but that the two pregnancy costs EX and FX are variable. EX is variable expenditures per pregnancy which the family controls to optimize pregnancy outcome given resource constraints. FX is the fixed cost of a pregnancy and although it may vary among families (just as RE may) it is assumed not to subject to choice for a specific family. As an example of FX, DeTray (1972) ha pointed out that pregnancy is not costless because there is always some, perhaps miniscule, risk of maternal mortality. Additional and more substantial costs would be associated with the time lost by the mother either simply during the confinement surrounding the birth or in a broader sense resulting from her decreased productivity either in the home or market because of the physical stress of pregnancy on her body. For example, she will probably require more rest, do certain tasks more slowly and be unable to perform certain very strenuous tasks particularly during the latter months of the pregnancy period. If such fixed costs result primarily from net time loss to the mother we would not be surprised to find that the fixed cost of pregnancy was higher for mothers whose price of time was higher.

Using this notation we can partition child associated costs as follows:

$$\text{Total Cost of Pregnancies} = n(EX + FX) \tag{3}$$

$$\text{Cost of Surviving Children} = p \cdot n \cdot K \tag{4}$$

We can therefore write the income constraint as

$$Y = S + n(EX + FX) + pnK \tag{5}$$

Substituting the expressions for C and S from (2) and (5) into (1) yields

$$U^{*} = U(P(RE,EX)n,\ Y - n(EX + FX) + pnK) \tag{6}$$

Taking first partial derivatives of this utility function with respect to EX and n the two variables that the family can control yields the following condition for utility maximization:

$$\frac{\partial U^{*}}{\partial EX} = U_c n \frac{dp}{dEX} - U_s(n(1 + \frac{dp}{dEX}K)) \tag{7}$$

$$\therefore \quad \frac{dp}{dEX} = \frac{U_s}{U_c - U_s K} \tag{8}$$

where $\frac{\partial U}{\partial C} = U_c$ and $\frac{\partial U}{\partial S} = U_s$, and

$$\frac{\partial U^{*}}{\partial n} = U_c \frac{\partial C}{\partial n} + U_s[(EX + FX) + pK] \tag{9}$$

$$\frac{U_s}{U_c - U_s K} = \frac{p}{(EX+FX)} \tag{10}$$

Note that the right hand sides of both (8) and (10) are identical being the ratio of the marginal utility foregone from consumption of S (U_s) to the net marginal utility of a surviving infant ($U_c - U_s K$).

Combining (8) and (10) yields the first order condition for a maximum that

$$\frac{\partial p}{\partial EX} = \frac{p}{(EX+FX)} \tag{11}$$

This implies that utility is maximized by equating the marginal product (and hence cost) of the survival function ($\frac{\partial p}{\partial EX}$) with the average cost of survival. Expenditures beyond this point will cause the family to allocate too much of S to the production of children; this is particularly true as children may be produced more economically by increasing n, the number of pregnancies.

We can further illustrate the trade off between increasing expenditures on survival with increasing the number of pregnancies by examining the conditions for producing a given $c = c_o$ at minimum cost. As above, let Total Child Cost = n(EX+FX)+pnK and form the Lagrangian,

$$W = [n(EX + FX) + pnK] + \lambda(c_o - pn). \qquad (12)$$

Now minimize with respect to n and EX yielding

$$\frac{\partial W}{\partial n} = (EX + FX + pK) - \lambda p) \qquad (13)$$

$$\therefore \quad \lambda = \frac{(EX+FX)+pK}{p} \qquad (14)$$

and,

$$\frac{\partial W}{\partial EX} = ((n + \frac{\partial p}{\partial EX} \; nK) - \lambda n \frac{\partial p}{\partial EX}) \qquad (15)$$

$$\therefore \lambda = \frac{1 + \frac{\partial p}{\partial EX} K}{\frac{\partial p}{\partial EX}} \qquad (16)$$

Combining (14) and (16) and solving for $\frac{\partial p}{\partial EX}$ yields,

$$\frac{\partial p}{\partial EX} = \frac{p}{(EX+FX)} \qquad (11a)$$

as above. It is important to note that while the couple is assumed free to vary either EX or n in this model to produce optimal c, in fact, the optimal expenditure per pregnancy is independent of n and is solely a property of the production function p(RE,EX) and perhaps the level of RE. While one might want to restrict the p function in several ways, the only requirement for a stable equilibrium is that the average product of EX has to be greater than the marginal product (ie., AP = p/FX, if $\beta = \frac{EX}{EX+FX} < 1$, the share of variable expenditure in the total cost of a pregnancy then (11a) requires that $\frac{\partial p}{\partial EX} = \beta \cdot AP$ for cost minimization). This further implies that the production function has to allow for declining average product at the point of equilibrium. While this doesn't imply that marginal product need decline, the fact that p has an upper bound of 1 (certainly) would seem to imply that marginal product should also be declining for large enough values of EX.

It should also be noted that the symmetry between cost minimization and utility maximization goes beyond the finding that the marginal conditions (11) and (11a) are identical. The complete generality of c_0 in (12) implies that the resulting minimum cost solution is the same for any desired c and suggests the following scenario for determining the desired number of children, the desired number of pregnancies, and expenditure per pregnancy. Using the marginal condition (11a) and knowledge of P(RE,EX) and RE, one could solve for optimal p, p*, and associated optimal expenditure per pregnancy, EX*, together these parameters determine the total price of a surviving child,

$$\text{Total Price per Survivor} = \frac{(EX^* + FX) + p^* \pi k}{p^*} \qquad (17)$$

Using this price, the families initial endowment and their utility function one could determine the desired number of children, C* and the desired number of pregnancies $n^* = \frac{C^*}{p^*}$. Of particular interest would be how these results would be affected by changes in the exogenous variables, Y, FX, RE and changes in the production function P(EX,RE). Note, for example, that if RE is a neutral shift parameter so that p=RE·f(EX) then the marginal condition (11a) is independent of the level of RE and so therefore is EX*. Thus for any given level of RE, p* would be determined and so would C* and n^* as above, however by definition $\frac{\partial p^*}{\partial RE} > 0$ so that $\frac{\partial c^*}{\partial RE}$ should also be positive because the price of a surviving child declines as RE increases, the effect on n^* is however undetermined and depends on the trade-off between the increase in C^* and the fact that given the increase in p* fewer pregnancies per survivor are required.

The Production Relationship

Having demonstrated the importance of the marginal condition (11a) in determining expenditures on pregnancy and pregnancy outcomes and the symmetry between the cost minimization and utility maximization approach, we shall now concentrate on a number of exercises in comparative statistics in an effort to determine the effect of changes in the exogenous variables on pregnancy expenditures.

In Figures 1 and 2 are shown graphically the equilibrium condition implied in (11a). In both figures expenditures (EX) are plotted on the X-axis and average product (AP), marginal product (MP) and βAP (average cost) are plotted on the Y-axis. In both figures $\frac{\partial p}{\partial EX}=MP > 0$ but declining $\left(\frac{\partial^2 p}{\partial EX^2} < 0\right)$, this is consistent with the notion that since p approaches 1 as an upper bound marginal product is likely to fall. Figure 1 illustrates the standard text book relationship between average and marginal products: summarized mathematically as -

$$\frac{\partial AP}{\partial EX} = \frac{1}{EX} \left(\frac{\partial p}{\partial EX} - \frac{p}{EX}\right) \qquad (18)$$

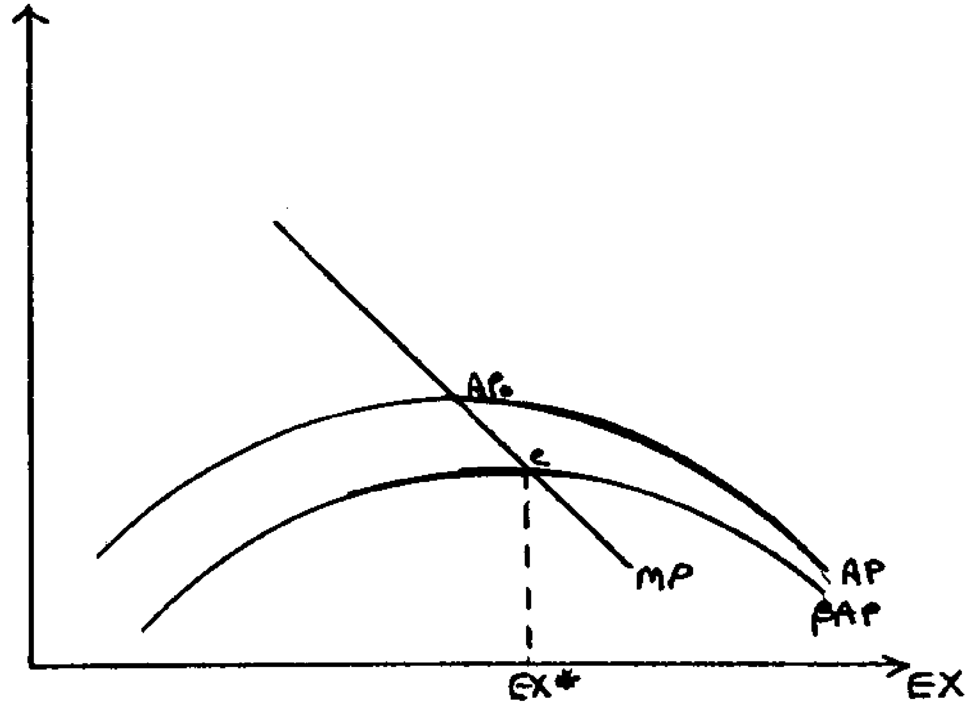


Figure 1 Equilibrium Conditions, U-shaped Average Product

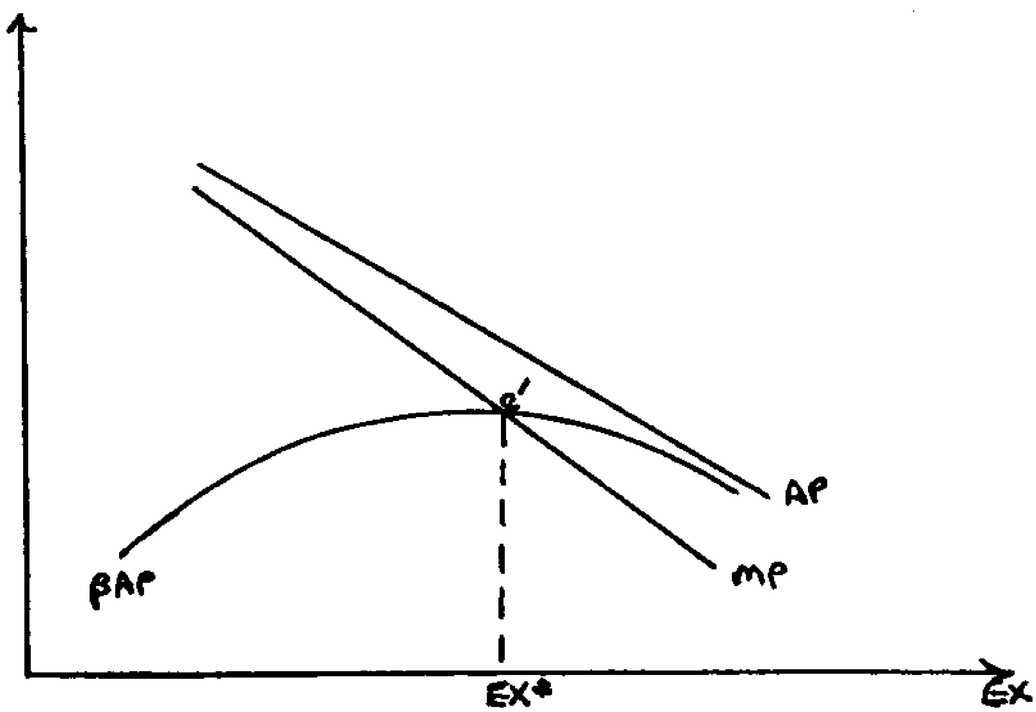


Figure 2 Equilibrium Conditions, Constantly Falling Average Product

where it is assumed that for small values of EX, MP $>$ AP so that AP rises, peaks where AP=MP and falls as MP dips below AP. Also shown is the function β AP where $\beta = \frac{EX}{FX+EX} < 1$ as above. Note that this curve has the same general shape as the AP curve but always lies below it and that $\beta AP \to AP$ as $EX \to \infty$. The intersection of βAP and MP at e is at the maximum point on the βAP function and defines EX*, optimal expenditure on pregnancy at a point where AP has already begun to decline (as discussed above).

Figure 2 is less restrictive but illustrates a similar relationship between β AP and EX. Here we only use the necessary condition that AP be falling at the equilibrium point and the implication that AP $>$ MP. As shown on the figure, AP $>$ MP for all values of EX so that both AP and MP fall as EX increases. This is not true for βAP however. Note that -

$$\frac{\partial \beta AP}{\partial EX} = \frac{1}{EX+FX} \left(\frac{\partial p}{\partial EX} - \frac{p}{EX+FX} \right) \tag{19}$$

and that (19) taken in conjunction with the marginal condition (11a) implies that β AP $>$ MP for small values of EX rises to its maximum at β AP=MP, point e', and then falls as MP $< \beta AP^*$. Note that as $EX \to \infty$, $\beta AP \to$ AP so that beyond e', β AP lies between MP and AP. Once again the point e', the point of intersection, defines EX* and its associated parameters p*, C*, and n^* under ceteris paribus conditions. Clearly then as seen in Figure 2, we do not require the more restrictive inverted U-shaped AP function to obtain a βAP curve properly shaped to yield the equilibrium condition.

Income Effect

Since the circumstances graphed in Figure 2 are sufficient for an equilibrium and somewhat more general than the conditions in Figure 1, I

---

*Note that $\left(\frac{\partial p}{\partial EX} - \frac{p}{EX}\right) < \left(\frac{\partial p}{\partial EX} - \frac{p}{EX+FX}\right)$ so that (19) may be positive although (18) is negative - this would be particularly likely where EX was small relative to FX.

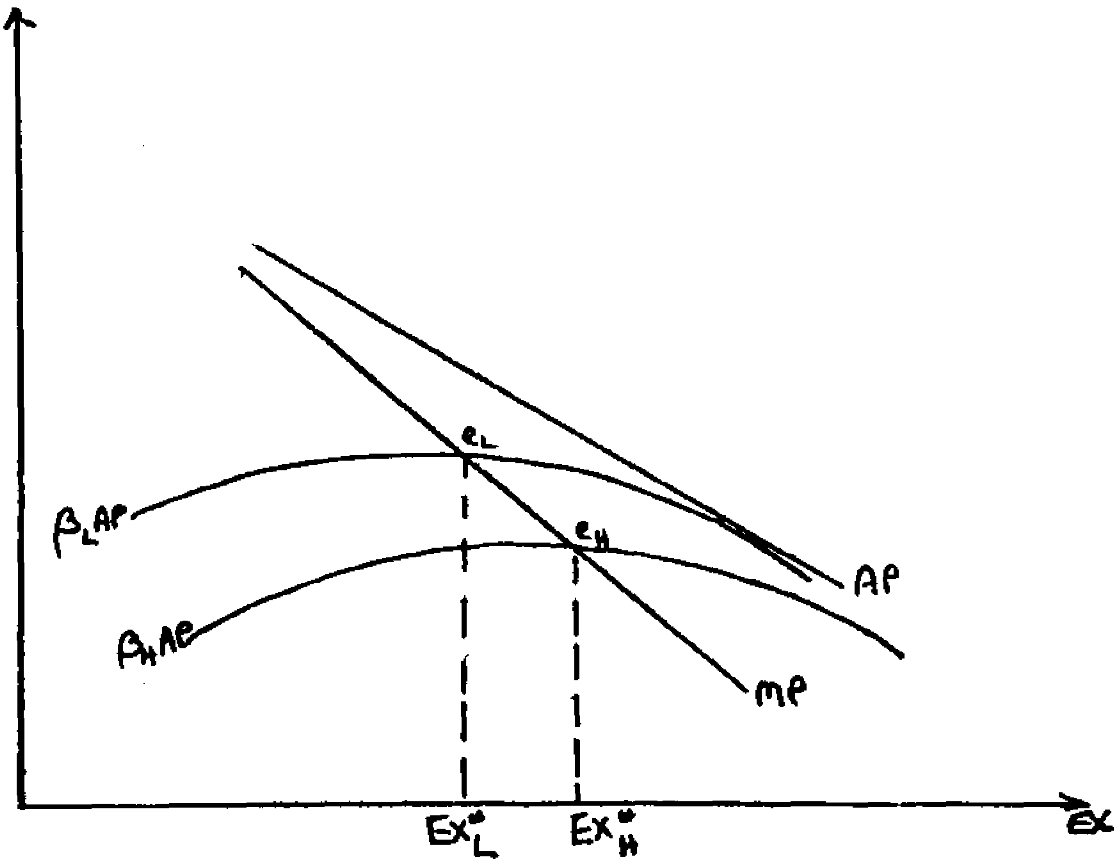


Figure 3 (II) Income Effect on Demand for Inputs

shall use variations of this figure to derive relationship between expenditures and changes in other exogenous variables. Consider first the effect of a change in income, Y on EX^*. If children are a normal good, $\frac{\partial C}{\partial Y} > 0$, and we would expect the desired number of children, C^*, to increase as income rose. Unless Y enters the calculation of EX^*, p^* will not change and the only way that higher income families can increase family size is to increase the desired number of pregnancies n^* until a new equilibrium between S and C^* is reached. One way around this dilemma is to assume that $\frac{\partial p}{\partial Y} > 0$, after all p is a measure of child health and child health as an aspect of child quality might well be related to income as is the quality of other goods (Grossman, 1973). This argument, while perfectly reasonable, is an unnecessary assumption in our relatively simple model and reduces the power of the symmetrical production/utility maximization relationship which otherwise greatly simplifies the model.

An alternative reason for expecting EX^* to be positively related to income can be demonstrated in Figure 3 and involves the consideration of the relationship between income and FX, fixed cost. It was argued above that the primary component of fixed cost was the loss of mother's time due to pregnancy. There is of course the actual time lost during the birth and maternal recovery period; however, generally during the prenatal period, particularly late in pregnancy, even healthy mothers will find that they tire easily, may require more sleep and may find certain tasks particulary those requiring strength or agility difficult if not impossible to perform. Such a reduction in the effective productivity of maternal time, as well as time actually lost, may be regarded as the fixed time cost of a pregnancy. It is important to distinguish this time cost from the variable time cost which may be associated with visits for prenatal care or time spent in classes in preparation for child birth and child care. Such time expenditures are variable and should properly be included in the expenditures (EX) category of pregnancy costs. Even if the properly defined fixed amount of time lost is the same for women of different income levels the cost associated with this time loss will not be. Standard results from household production theory imply that the cost of time is positively correlated with income, and, hence one would expect fixed costs FX, ceterus paribus, to be positivel correlated with income and β (the share of variable cost in total cost) will be less at each level of EX as income rises. The implications of this line reasoning are demonstrated in Figure 3. AP and MP are as before, the time loss associated with pregnancy are assumed the same for females at high (Y_H) and low (Y_L) income levels; however, fixed cost for higher income mothers is higher than for low and correspondingly $\beta_H < \beta_L$ (where β_H is β for high income mothers, and β_L is the B for low income mothers). Therefore, β_L AP lies above β_HAP and intersects MP at e_L to the left of e_H the intersection point for β_HAP, the implication is that $EX^*_L < EX^*_H$ and that expenditures on pregnancy will be positively correlated with income.

This not unexpected prediction is a testable hypothesis generated by the model rather than being an assumption as in other studies. Morever, the model implies that marginal and average product will negatively correlate with income.

The Production Function and Efficiency Effects

Having examined the expected effect of changes in income on pregnancy expenditures I would now like to examine the effect of differences in RE (exogenously determined reproductive efficiency) and, differences in the production function itself on expenditures. Thus far, we have characterized the production function as having a positive but falling marginal product, falling average product around the equilibrium point and being asymptotic to p=1 as EX becomes large. Such a function is shown in Figure 4 and labeled P_o. Note the intercept at a implies that even if EX were zero there would still be some possibility of survival, this seems quite likely given the positive albeit low infant survival rates observed in extremely poor countries and of course among primitive populations.

The existence of this positive intercept points out some interesting problems in defining differences in efficiency among different producers. Efficiency strictly defined refers to the ratio of total output to total input. It is not unusual to observe different firms in the same industry producing at different levels of efficiency, with different factor proportions and different levels of output. Differences in efficiency may be attributed among other things to differences in technologies employed (particularly vintage effects), economies or diseconomies of scale or differences in the level of some otherwise undefined input usually called entrepreneurial ability. The usual assumption is that differences in efficiency result from differences in the marginal products of factor inputs - more efficnent firms use factors better hence output is higher and so are the marginal products of particular inputs at a given level of production. Recently, this notion of differences in efficiency in production has been applied to the household production model of consumer demand. So called "environmental" variables, of which education is the most widely studied, have been thought to affect the production of commodities in the home by changing the marginal product of inputs in the production of household commodities. Micheal (1972) has tested the hypothesis that such an increase in efficiency in the household acts as though it augmented family income, increasing the consumption of luxuries and decreasing the consumption of necessities among those families with more education at a given income level. Grossman (1972) and Inman (1976) have extended this notion of increased efficiency resulting from higher levels of education to the production of health.

Notice in Figure 4, that increases in the marginal product of variable inputs are not the only source of increased efficiency, particularly in the production situation we are considering. Consider a mother whose production function is represented by P_1. P_1 is derived from P_o by merely shifting the function upward by the amount b-a the differences in the level of output with zero variable inputs. The marginal product of EX is the same for all levels of EX below the level where p approaches 1 asymptotically, EX'. For levels of EX below EX' the producer on P_1 is more efficient than the producer using P_o although the marginal product of the input is the same.

Consider now P_2, it has the same intercept as P_o but a higher marginal product. In a sense, the advantage of P_2 relative to P_o represents the advantage accruing to the usually examined environmental variables like entrepreneurial ability or education. Moreover, it is possible to think of intermediate production functions such as producers with lower intercepts but higher marginal products - they may be less efficient for some lower levels of EX but more efficient at higher levels. Production functions such as those graphed in Figure 4 are not widely used in economics particularly since situations with positive output with no endogenous inputs or limitations on output are rarely encountered. They do have relevance to the study of production of human capital type attributes in individuals since both the genetic endowment (a possible

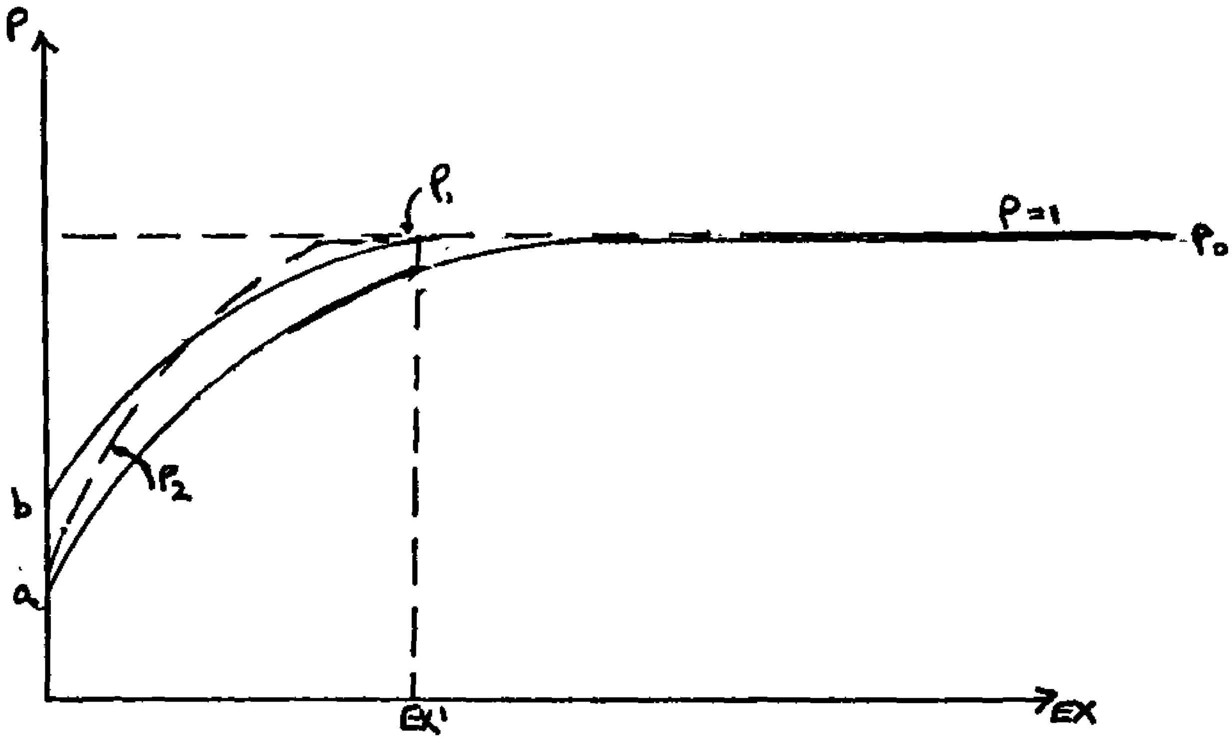


Figure 4 Survival Production Functions

variant of the intercept) and asymptotically limited output may be important in the production of attributes like survival or particular skills. Grossman (1972) uses an asymptotically contained production relationship to define the production of healthy time from the stock of health capital (endogenous variable) in a given time period (eg., one can be healthy no more than 7 days a week or 365 days a year).

I now would like to consider the probable effects of changes in the conditions of production on the marginal conditions and the endogenously determined EX*, p*, etc. As an example, consider the function:

$$P = 1 - ae^{-bEX} \tag{20}$$

where p is probability of survival, EX expenditures and a ($a < 1$) and b positive coefficients which will reflect the different notions of efficiency which I have discussed. Note that (20) has all the desirable attributes we have used earlier,

$$\frac{\partial p}{\partial EX} = abe^{-bEX} > 0 \tag{21}$$

$$\frac{\partial^2 p}{\partial EX^2} = -ab^2e^{-bEX} < 0 \tag{22}$$

and

$$\lim_{EX \to \infty} p = 1 \tag{23}$$

Note, further than (1-a) is the level of p when EX=0. Moreover,

$$AP = \frac{1-ae^{-bEX}}{EX} \tag{24}$$

may either rise and fall with increasing EX (as in Figure 1) or fall for all levels of Ex (as in Figure 2) depending on the values of the parameters of a and

Referring to Figure 5, consider the effect of a change in a on EX*. The curves MP_o and βAP_o are presented as the initial condition. They intersect at e_o and the resulting optimal expenditure is EX_o. Now suppose

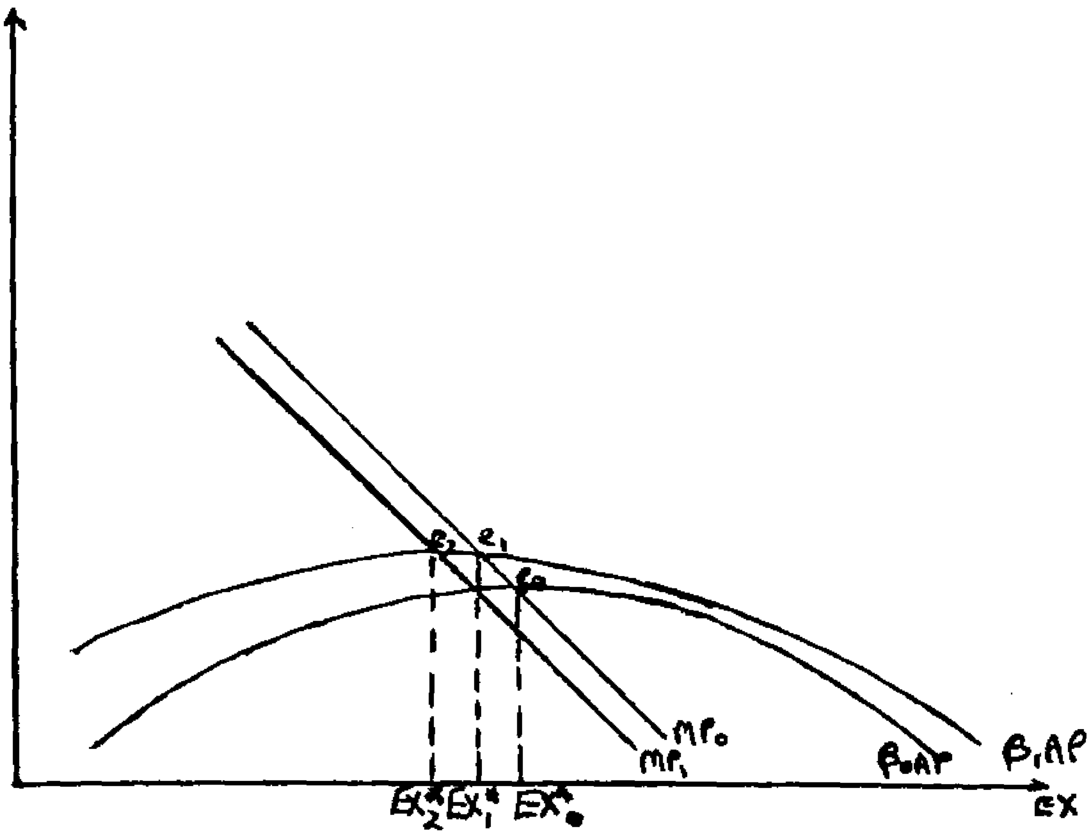


Figure 5 Effect of a change in $\underline{a}$ on EX^*

that a decreases (note that a decrease in a increases the intercept (1-a) and shifts the p function upward). A decrease in a will cause the βAP curve,

$$\beta AP = \frac{1 - ae^{-bEX}}{EX+FX} \tag{25}$$

to shift upward to βAP_1. If MP_o is not affected the new point of intersection will be at e_1 and the optimal $EX^*_1 < EX^*_o$. In other words, a plain shift upward in the production function which doesn't change the MP over the relevant range will cause expenditures to decline because the same level of p^*, the desired level, can be had more cheaply.

Notice with the production function given in (20) marginal product (21) is also a function of a, in fact

$$\frac{\partial MP}{\partial a} = be^{-bEX} > 0 \tag{26}$$

This implies that with an decrease in a, MP will fall to a new level MP_1 and the new intersection , e_2 where $MP_1 = \beta AP_1$ will be at an even lower level of EX, EX^*_2.

We now turn to an analysis of the effect of changes in the slope of p only, that is changes in b. Referring to Figure 6, once again AP_o, MP_o, e_o and EX^*_o reflect the initial conditions. Note:

$$\frac{\partial MP}{\partial b} = ae^{-bEX}(1-bEX) \tag{27}$$

so that the effect of a change in b on MP is ambiguous and depends on the sign of (1-bEX).* The problem of determining the effect of a

---

*Of course, ae^{-bEX} is always positive.

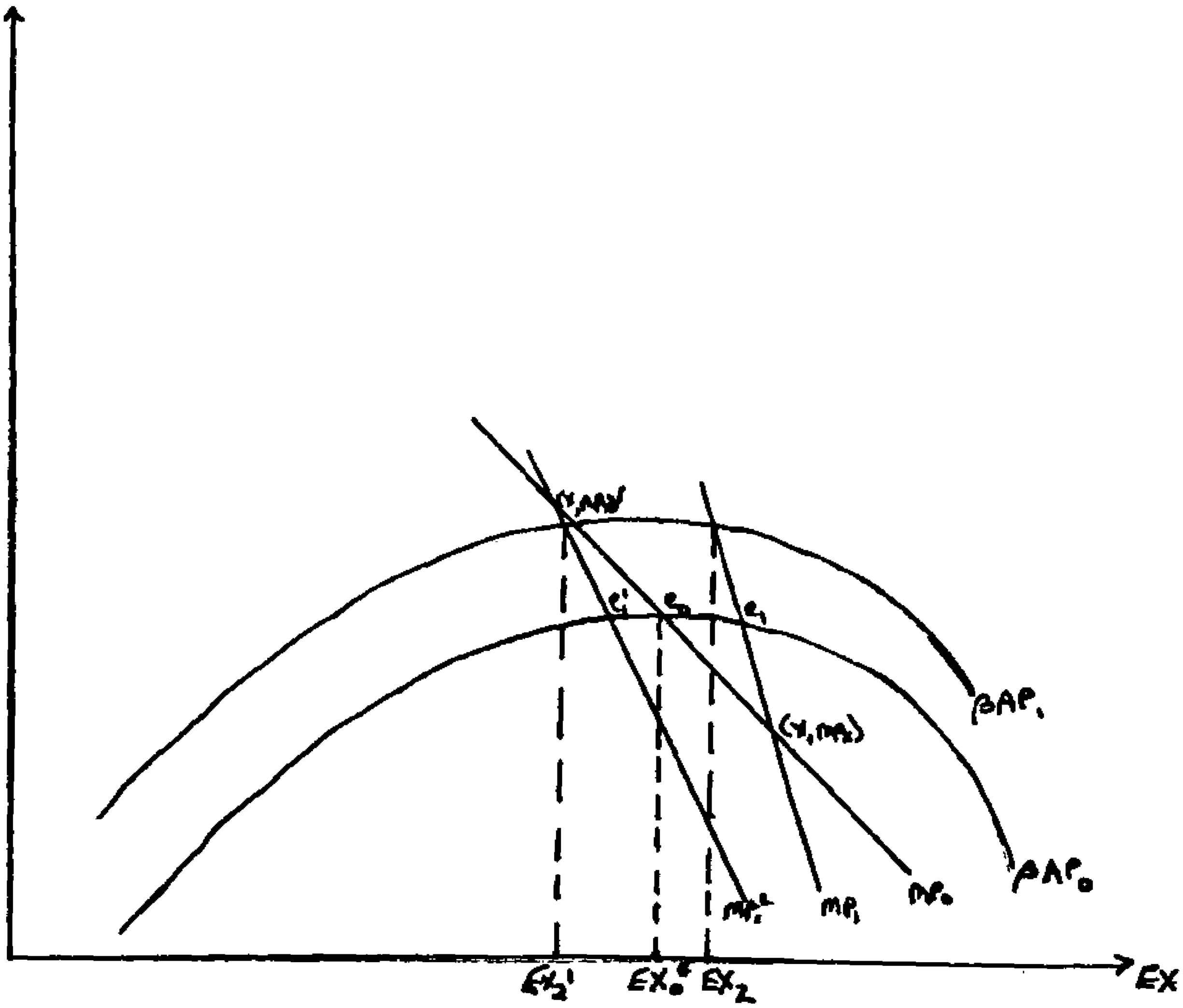


Figure 6 (III). Effect of a Change in <u>b</u> on EX^d

change in b is illustrated in Figure 6. For some value of EX, call it x, (1-bx) = 0 so that $\frac{\partial MP}{\partial b} = 0$. For values of EX less than x, $\frac{\partial MP}{\partial b}$ is positive and for values of EX greater than x, $\frac{\partial MP}{\partial b}$ is negative, hence the MP curve may be said to pivot through the point (X, MP_x) to some new position as b changes. If $x > EX^*_o$, then for example after an increase in b the MP curve will rotate clockwise around x to a new position at MP_1 and the new equilibrium point along βAP_o will be at e_1 and expenditures will increase. However, if $x < EX^*_o$, then the new MP, MP_1', curve will lie below the original curve MP at the point where it intersects βAP_o and the new point of intersection with βAP_o, e_1', will be at a lower level of expenditures.

The prediction of the final equilibrium position is further complicated by the shift induced in βAP by a change in b. Since

$$\frac{\partial \beta AP}{\partial b} = \beta a e^{-bx} > 0 , \tag{28}$$

an increase in b will cause βAP to shift upward to βAP_1. An upward shift in βAP will always tend to reduce expenditures along a given MP curve. Thus in the case where $x < EX^*_o$, the new intersection between βAP, and MP_1' will be at an even lower level of expenditures EX_2'. Moreover, in the case where $x > EX^*_o$, the result is truely ambiguous as the shift in βAP to βAP_1 will tend to reduce expenditures while the shift in MP to MP_1 will tend to raise expenditures. The new equilibrium point, e_2, illustrated in the figure, where EX_2 is greater than EX^*_o but less than EX_1 is only one of three possible outcomes which include $EX_2 = EX_o$ or $EX_2 < EX_o$. Although the succeeding exercise demonstrates that one can enumerate more possible patterns of functional shifting that are consistent with a reduction in expenditures, the actual change in EX induced by a change in b depends on the actual parameter values of a and b and therefore it is not correct to infer that a reduction in expenditures is the "more likely" result.*

How may we interpret these results in light of the earlier discussion of efficiency and particularly our concern about shifts in RE. It is tempting to interpret shifts in a (or only the intercept term) as being analogous to shifts in reproductive efficiency. This is particularly true because changes in a may affect p independently of the level of expenditure. Thus a has many aspects of a pure endowment effect that might be related to the genetic or biological production state of the mother. With the functional form used as an example above, a shift in a which raises the intercept will cause marginal product to decline - perhaps a result not consistent with everyone's notion of reproductive efficiency; however, given the asymptotic nature of any production function for p it would seem almost inevitable that an increase in the intercept

---

*Michael Grossman has pointed out to me that one can obtain the same result by implicit differentiation of the optimal condition (11). Thus, utility is maximized when

$$\frac{\partial p}{\partial EX} = \frac{p}{(EX+FX)} = \beta AP \qquad (11)$$

and substituting for $\frac{\partial p}{\partial FX}$ from (21) and for βAP from (25) yields

$$abe^{-bEX} = \frac{1-ae^{-bEX}}{EX+FX} \qquad (28a)$$

or

$$1 = ae^{-bEX}(1+EXb+FXb) \qquad (28b)$$

If we implicitly differentiate (28b) with respect to b remembering that EX will change as b changes to maintain the optimal point, we get

$$\frac{dEX}{db} = \frac{FX-FX\cdot EXb-EX^2b}{(EX+FX)b^2} \qquad (28c)$$

Since the denominator of (28c) is always positive, the sign of $\frac{dEX}{db}$ will depend on the sign of the numerator and will be positive, negative or zero, depending on whether

$$FX \lesseqgtr EXb(FX+EX) \qquad (28d)$$

respectively. Thus, as above, the sign of $\frac{dEX}{db}$ is indeterminate a priori. Moreover, as above, with FX predetermined, the sign depends on the optimal value of EX that corresponds to a given b and hence on the actual parameters of the function itself.

would lead to a decrease in the slope as p approached 1.

It is tempting to view b as a measure of an efficiency effect associated with an environmental variable such as education. This is because changes in b affect only the marginal product of the inputs rather than the intercept. Also note that,

$$\frac{\partial p}{\partial a} = -e^{-bEX} < 0 \qquad (28)$$

so that the impact of changes in a on p depends on the value of b. If a is regarded loosely as RE, an exogenously given input of biological factors then the size of the parameter b will determine the marginal product and mix of inputs in a manner analogous to the general environmental variable discussed in the literature.

We may conclude with the following predictions from the model:

(1) a change in RE, which is essentially incorporated in shifts in the entire production function, will tend to be negatively correlated with changes in EX^* (note, however, that since the total price per survivor is also negatively correlated with RE, C^* should be positively correlated with RE as children become cheaper to produce);

(2) changes in environmental variables such as education, which change only the marginal product of inputs will have an ambiguous effect on EX^*. Changes associated with decline in MP should cause expenditures to fall, however, if MP rises, the sign of the resulting change in expenditures is ambiguous. (Note, however, that if upward shifts in MP are to be regarded as representing increased efficiency, the total price per survivor should fall under these circumstances, although EX^* may increase, this should also lead to an increase in C^*.)

Value of Experience

Thus far in our discussion we have treated the decision making process as though it were beirg determined in a certain world with complete informatio[n]. In fact, p is a probability and therefore outcomes are uncertain so that a natural extention of the model would be to expand it to include expected utility analysis and a discussion of decision making in this context. We shall not attempt this extension but rely on Ben-Porath's (1973) experience in this area that such extensions are not likely to lead to significantly more complete models without substantial additional restrictions being placed on the utility function of the individual family.

Moreover, such an extension is not likely to shed much light on the significant question as to how expectations are formed and modified during the family formation process. It is worth noting that while we have assumed a one period model with decisions on critical variables being made at the beginning of the family formation period, in fact, family formation takes place in a substantial time continum with one pregnancy following another in sequence. During this sequential process many variables can change, expectations about income, attitudes toward and the demand for children and expectations about the production function for children p(RE,EX). In particular, in the terminology of the previous section, families' expectations about both RE and $\frac{\partial p}{\partial EX}$ might well change leading them to revise their decisions about EX^*, p^*, C^* and n^*.

Substantial work in this area dealing with the effect of exogenous child attributes on fertility decisions has been reported in a series of papers by Welch and Ben-Porath. Although they have been primarily concerned with the effect of the sex mix of surviving children on family size they have indicated that an extension of their model to account for other exogenously determined attributes such as mortality is possible. The extension of such a model to measure empirically the effect of pregnancy losses on the decision to have an additional child has been investigated by Williams (1976).

Welch (1974) provides a mathematically detailed sequential decision model of the effect of prior experience with the sex of offspring on the decision to have additional children. His conclusion based on what he feels are reasonable estimates of the relevant parameters is that the effects of learning (i.e., changing expectations about the sex ratio) are not dominant. Williams (1976) applies Welch's derivation to the case of completely exogenously determined mortality but without definite predictions about measurable parameters. Neither consider the situation which has been our concern so far, that is, the effect of experience on EX* - endogenously determined expenditures during the current and future pregnancies.

Because Welch does provide an argument about the expected effects of experience it is worth calling attention to empirical differences in the determination of the sex ratio and infant loss rate. For one thing, I have argued that the infant loss rate is partially endogenously determined by family expenditures on pregnancy - historically the sex of the unborn child has not been subject to parental control. Secondly, the expected values of the sex ratio and infant loss rate within the population are of a different order of magnitude particularly in developed countries: Welch estimates the sex ratio at about .51 male while recent U.S. data indicate that survival of offspring during late pregnancy (over 20 weeks gestation) and infancy (less than 1 year) exceeds .95. Thus, if the family plans 3 pregnancies the probability that all three will be boys is .125 while the probability that all three will survive is .857, seven times greater. Moreover, if there is some desire for balance, as Welch assumes, the probability in 3 pregnancies of a mix of sexes is .75 while the probability of a single loss only is .135.* If w adopt the information theory approach to the value of information (i.e., that the more unlikely an event is the more information its occurrence contains), then clearly there is more potential information to be gained about the potential survival of progeny from family experience in this area than there is to be learnt about potential differences in the sex ratio.

---

*Moreover, it is unlikely that there is any desire for "balance" between dead and surviving children as there might be between boys and girls.

Let us now consider the implications of gaining experience about child survival patterns as children are born. Recall that the original model we considered was a one period decision model where decisions were made about EX^*, p^*, c^* and n^* at the beginning of the child bearing period and expected to be held throughout the period of family formation. Since p^* is a probability < 1, at the end of n^* pregnancies there will be some distribution of surviving children among families whose initial goals were all the same. Some families will have achieved C^* children, these we will call the lucky ones, others fewer children with the most unlucky families having no surviving children. Following Welch, we shall discuss the results of such a distribution of outcomes as encompassing an income and a learning (price) effect.

Consider the family for whom after n^* pregnancies, $C = 0$, they have experienced an income loss equal to $n^*(EX^*+FX)$. And if there is no learning (reformulation of expectations) or disutility associated with pregnancy losses, they will be in the same situation as a family starting at that lower income level. In this rather extreme example, they optimize again and continue attempting to have children with a new $C^{*'}$ and $EX^{*'}$ appropriate to their new lower income level. Their loss in the production of children will not be absorbed totally in the demand for children but rather spread throughout all consumption. Note, however, that if C is a normal good (i.e. $\frac{\partial c^*}{\partial y} > 0$) $C^{*'} < C^*$ and if the income effect on EX^* is as we derived earlier $EX^{*'} < EX^*$. It does not seem unreasonable to expect similar behavior from all families whose attained C is less than C^* - this in fact is a form of the replacement behavior we discussed earlier.*

Consider now the learning or informational effect of having experienced n^* pregnancies with something other than C^* success. Clearly we are dealing with a second type of uncertainty here - that is we have assumed up to this point that the production function, $p = f(RE, EX)$ and the level of RE were known with certainty and that the only uncertainty of outcome resulted from p^* being less than 1. In the real world it is very unlikely that either RE or $f(RE,EX)$ will be known with certainty and the household can only form estimates as to their values. The parameters of interest to the household are RE and $\frac{\partial p}{\partial EX}$ and we shall denote their estimate of these values by $\widetilde{RE}$ and $\widetilde{MP}$. Fortunately, our earlier discussion of the symmetry between utility maximization and cost minimization as well as our exercises in comparative statistics regarding these variables will stand us in good stead in this area.

First, we should distinguish between losses which convey information about the production process and losses which do not. Pregnancy losses due to apparently random events outside of the pregnancy process itself such as accidents or deaths by violence presumably convey little information about expected outcomes resulting from subsequent pregnancies. Their influence on subsequent pregnancy related decisions will be felt only as the result of a pure income loss. On the other hand, during a pregnancy specific information about the pregnancy process and the value of

*Notice, however, that because pregnancies and children occur in discrete rather than continuous bundles this effect may be muted as C attained approaches C^*. For example, if $p^* = .9$ and $C^* = 3$ then 3.33 pregnancies will be required on average to attain C^*. The family may estimate n^* at 3-4 pregnancies judging itself more luck if it takes only 3 and less lucky but still not unfortunate if it takes 4. After completing 3 pregnancies with 2 survivors it may well continue on to its fourth pregnancy under its original scenario without revising any of its initial targets.

inputs and expected outcomes may be gained even if the resulting infant survives (e.g., information about potential Rh sensitization resulting from pregnancy is usually obtained as a result of specific blood typing of parents during the first pregnancy). Let us consider the effect of such information on subsequent pregnancy decisions whether or not an unexpected loss has occurred.

Let us once again consider the case of the household with no success after n^* pregnancies. If they do not reevaluate $\widetilde{RE}$ and $\widetilde{MP}$ then they will reformulate a new decision plan using these same variables and only the income effect analyzed above will cause a change in $EX^{*'}$. However, if p^* was thought to be high it is unlikely that they will not be tempted to revise their estimate of p^* which depended on their estimate of $p = f(EX,RE)$. They may revise their estimate $\widetilde{RE}$ which we have loosely defined as being associated with the height of the function or their estimate $\widetilde{MP}$.*

Let us briefly recall the results of the previous section where we examined the effects of changes in RE and MP using the function $p = 1 - ae^{-bEX}$ as an example of the type of functional relationship which might well relate RE and EX to p. Recall that a change in A which we felt could be interpreted as being related to biological efficiency (RE) was seen to be negatively correlated with EX^* and that this effect was only reinforced because of related induced shifts with this function in MP. Recall also that shifts in b which we interpreted as only effecting MP and being associated with efficiency in production associated with environmental

---

*We have previously noted that since $p = f(EX,RE)$ is generally asymptotic to 1, the height and the slope of p will probably be related for some large enough value of EX as p approaches the asymptote. However, we shall continue to analyze these effects separately.

variables were negatively correlated with $EX.^*$

Although we have assumed that the household views p as being assymptotic to 1 as $EX \to \infty$, it is conceivable that perceptions as to differences in reproductive efficiency - particularly biological processes may be associated with perceived differences in the level of the asymptote. For instance, if $p' = c - ae^{-bx}$ where $0 < c \leq 1$ then perceived variations in c would effect expectations as to the level of p attainable with a given quantity of EX as well as the level of p attainable as $EX \to \infty$. Moreover, c may be regarded as a pure shift parameter, i.e.

$$\frac{\partial p'}{\partial c} = 1$$

$$\frac{\partial p'}{\partial EX} = abe^{-bEX} \quad \text{so that MP is independent of c}$$

$$\text{and} \quad \frac{\partial AP}{\partial c} = \frac{1}{EX+FX} > 0 \ .$$

With these properties in mind and recalling the analysis of Figure 5, we should expect c to be negatively correlated with EX^*. Thus couples who view their bad luck as being evidence of a lower level production function due to biologically determined factors independent of the level of EX will spend more per pregnancy.

We may summarize the results of experience as follows:

(1) Couples who experience a level of $C < C^*$ with expenditure EX^* will experience an income loss effect which will tend to reduce $EX^{*'}$ and $C^{*'}$ the new targets during the second decision making period. Depending on the size of the income effect they may decide to continue $n^{*'} > n^*$ or cease having additional children.

(2) If couples further experience a learning effect which they interpret as reflecting a biological reproductive efficiency effect of the type characterized by changes in <u>a</u> or <u>c</u> as discussed above, they will tend to increase their target EX^* to $EX^{*'}$ to compensate for their poorer position. Notice that such movement is contrary to the income effect and the question of which effect dominates becomes largely a matter of empirical determination. However, a reduction in $\tilde{p}^*$ associated with such bad experience will tend to increase EX^* and this indicates that the household will view the price of children as having risen and this will retard fertility, (i.e. $c^{*'} < c^*$) in a manner reinforcing the income effect.

We may now relax one more element of this model and move towards a more realistic formulation of the fertility related decision making process. Initially, we assumed a one period model where households chose the utility maximizing values of EX, c, n, and p and stuck to that decision throughout the period of family formation. In this section we recognize that since p^* is likely to be less than 1, certain families will not attain c^* after n^* pregnancies even though they expend EX^* per pregnancy for these families regardless of any revisions in their expectations concerning p^*, it is logical to regard them as being in a situation similar to other families with a reduced income level and allow them the possibility of trying again (i.e., have pregnancies beyond n^*). Moreover, if they also revise their expectations regarding $p = f(RE,EX)$ this will also effect their decision regarding c^*, n^*, EX^* and p^*.

It is easy to see that we can further relax our assumptions about when decisions are made and consider a complete sequential decision making procedure of a special kind. For example, consider the situation cited earlier where $\tilde{p}^* = .9$ and $c^* = 3$ so that $n^* = 3.33$ or 3 to 4. After the first pregnancy most couples will experience a success but some will have had a failure. In particular, they now know that they will require at least 4 pregnancies to have children and if they regard the loss as indicating a lower than expected child production potential this may also effect their decisions about future expenditures on children as well as desired family size. Within the context of our one period model we may regard such couples as modifying their behavior to include possibly new values of EX^*, c^*, n^* and p^* as though these new values would apply to all <u>future</u> pregnancies. Moreover, it is not unlikely that if p^* is considerably lower than in our example or c^* considerably higher that couples who experience success may reevaluate their position in light of favorable results after each pregnancy. So long as households follow the procedure of minimizing the cost of producing the remaining children they desire then we need only be concerned with how their experience alters their view of the functions influencing the marginal condition (9a) in order to make predictions about EX^*. The stock of children already attained as well as the sequential path of pregnancies which lead to the current levels of c and S will influence decisions as to whether to have additional pregnancies but will not influence our predictions about the effect on expenditures per birth.

Empirically we may, however, encounter a problem in estimating the effect of past experience on expenditures on pregnancies in a cross section of pregnant women. This is a problem which has been encountered elsewhere in econometrics and is known generically as the censored sample problem. That is there will be a group of women whose experience with pregnancy will have been so good that they will reach desired family relatively quickly and drop out of the birth cohort for them we will not be able to measure the effect of their cumulative good fortune on EX for they will have ceased having children. On the other end of the spectrum are those whose experience has been so bad that due to both income and primarily price effects they decide to curtail fertility: for them too, we have no information about what they would have spent on an additional pregnancy as they have decided to have no more. The exact effect of this drop out phenomena on estimation biases is hard to predict a priori; however, it is worth pondering the implications of changes in outcomes on sequential fertility decisions. Billewicz (1973) using data from Scotland reports that women who have had a string of unsuccesful pregnancies tend to quit child bearing after their first success while women with a string of successes tend to quit after the first failure. Both forms of behavior would tend to indicate a significant learning response to bad outcomes which would be reflected in cessation of child bearing.

Wantedness

One of the critical assumptions of the model developed thus far and of most current fertility and child health models is the assumption that children are regarded by the household as a good - a wanted commodity. Yet there is substantial evidence that infanticide has been practiced in both past and contemporary societies. Moreover, it has been observed that infant mortality having reached a plateau in this country in the early 1950's did not begin to fall rapidly until the mid 1960's contemporaneously with the introduction of modern forms of birth control. The decline in infant mortality has continued and in fact, accelerated since the legalization of abortion in the early 1970's. This has lead some to argue that the degree of "wantedness" of an infant (pregnancy) may be a significant determinant of pregnancy outcomes.

A study by Pakter and Nelson (1974) stresses the importance of the relatively recent introduction of fertility control measures in shifting the distribution of births by maternal age and parity into the more favorable categories--although an explanation for why these categories have traditionally experienced low infant loss rates is not given. Fuchs (1974) also notes the importance of declines in the proportion of births of birth order four or above (where death rates have traditionally been greater) in explaining part of the rapid decline in U.S. infant mortality since 1965. He attributes this shift to a general improvement in birth control. He also credits the improvement in birth control technology with presumably decreasing the proportion of "unwanted" live births during this same period and states, "The infant mortality rate for 'unwanted' children is undoubtedly many times higher than for wanted children."

On the other hand, in discussing an interview study of mothers conducted immediately post-partum, Morris, Udry and Chase (1973) conclude "that the prevention of unwanted pregnancies will cause no practical reduction in low birth weight birth rates." Birth weight is an important index of child health at birth and an important predictor of infant mortality. In addition, a study of births to Swedish women (Hultin and Ottosson, 1971) who had requested and been denied abortions under a somewhat elective abortion control system showed no reduction in most indices of infant health at birth and no increase in the perinatal mortality rate among the group of presumed "unwanted" births as compared with a matched control group of births to women delivering in the same hospital.*

Unfortunately, a single definition of "unwantedness" has not been consistently applied in studies cited. Fuchs seems to imply, using aggregate data, that higher birth order pregnancies are more likely to be "unwanted" but does not elaborate the concept further.** Morris, Udry and Chase rely on the mother's response to direct question about whether the baby was wanted prior to conception only after a viable infant has been delivered. Hultin and Ottosson accept the request for abortion as prima facie evidence of "unwantedness."*** While all of these may be justifiable empirical instruments for "wantedness", of primary significance for our investigation is

---

*Hultin and Ottosson do report a significantly higher rate of malformations among the "unwanted" births but cannot explain the difference.

**A serious reservation about assuming that higher birth order births are in the aggregate less likely to be "wanted," results from the fact that nearly all illegitimate births occur as first or second births and presumably most of these are "unwanted."

***Moris, Udry and Chase point out that a problem with measuring the "unwantedness" of an individual birth may result from attempting to distinguish between not wanting the child ever and a mere timing error in the spacing of the birth. Although many fertility studies have assumed that the measure of unwanted fertility that produced the highest estimate was best, they argue that for purposes of evaluating the health of the infant the more restrictive definition is appropriate. In fact, one would want to examine the nature of the individually perceived timing error before making such judgement. Note, that in the Swedish study it isn't possible to distinguish requests for abortions that resulted from timing errors or from not ever wanting the child.

whether "wantedness" can be given a consistent economic definition and within this context what can we say about its importance in determining EX and pregnancy outcomes.

Fortunately, "wantedness" can be given a straightforward economic definition which is consistent with standard demand theory. Consider Figure 7, it represents the standard economic demand analysis applied to the demand for children. The price of a child is plotted on the vertical axis (all children have the same price) and the quantity of children on the horizontal. The demand relationship, D between price and quantity demanded is downward sloping to the right, not because of strict declining marginal utility of children but because with a fixed income the marginal rate of substitution for children in relation to other goods declines as the number of children increases. If the price of a child is Po then q_o children will be desired; once a family has q_o children any additional child becomes unwanted because it is too expensive. At a lower price more children will be wanted than q_o and at a higher price fewer than q_o children are wanted. Notice also that the net loss of an additional child beyond q_o is not the price P_o but only the area represented by the triang ABC in the figure which is the difference between what the family would have been willing to pay ACFE and the actual cost ABFE.

In the context of our model where the expenditure on children and hence the price is endogenously determined is somewhat different. Consider the optimal condition (9) which results from differentiating U^* with respect to n

$$\frac{\partial U^*}{\partial n} = p(U_c - U_s K) - U_s(EX+FX) \qquad (9a)$$

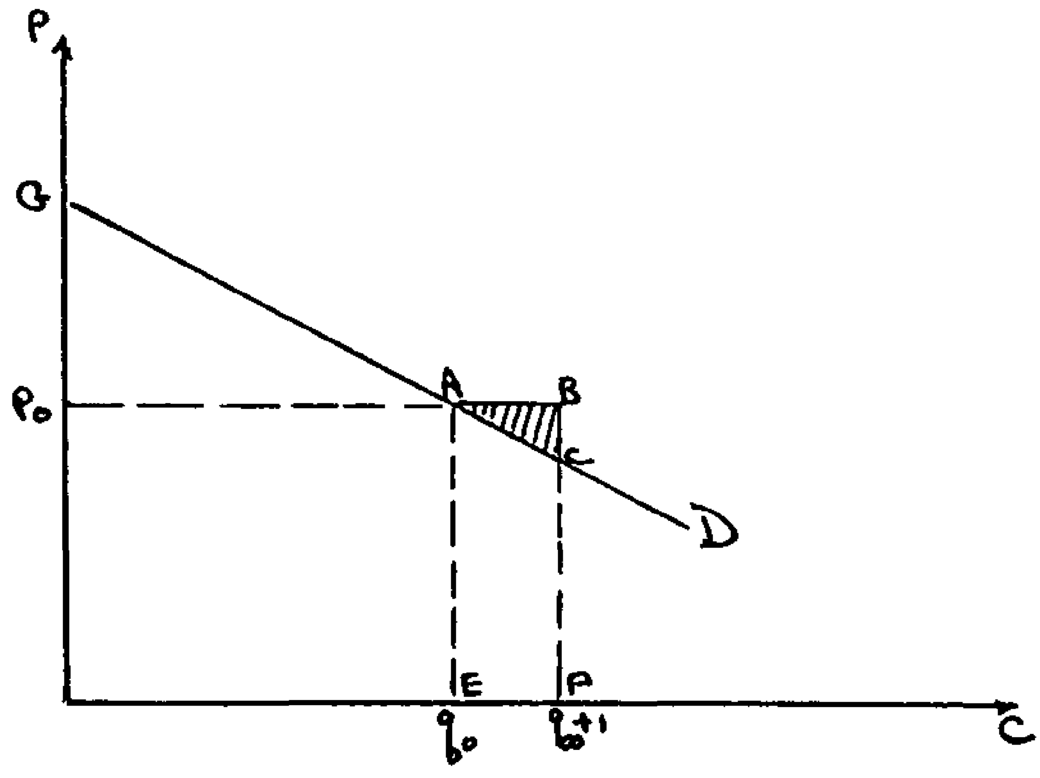


Figure 7 A Demand Function for Children

the first term represents the expected (probable) net gain in utility from an additional pregnancy, the second term the certain loss as a result of pregnancy related costs. If this derivative is greater than zero, an additional pregnancy will be "wanted" in the sense that the expected gain will exceed the certain loss. When (9a) equals 0 as in (9), we have reached the optimal number of pregnancies and when (9a) is negative an additional pregnancy (child) can be said to be unwanted. We may rewrite (9a) as

$$p\,U_c \lesseqgtr U_S(EX+FX+pK) \qquad (30)$$

and graph the functions and discover the optimal point (EX^*) by their intersection. Thus in Figure 8, EX is measured on the X-axis and the two marginal utilities (U_c, U_s) on the Y-axis. pU_c is plotted as an increasing function of EX, essentially the function f (RE,EX) multiplied by U_c which is assumed constant for any given value of c. $U_s(BX+FX+pK)$ is also an increasing function of EX and must lie above pU_c beyond the point of intersection EX^*_o for this optimal condition to represent a maximum.*

If EX^* defines optimal p^* then it also defines optimal c^* and n^*, however once we have reached c^* children or n^* pregnancies this intersection can no longer take place at EX^*. By definition of unwantedness for the

---

*Note that $\dfrac{\partial U_s(EX+FX+pk)}{\partial EX} = U_s - U_{ss}(1 + \dfrac{\partial p}{\partial EX}\ k)$

since S is a good $U_s > 0$ and in a two good world one would expect convexity of the indifference function to require $U_{ss} < 0$ so that $U_s(EX+FX+pk)$ will be positively sloped.

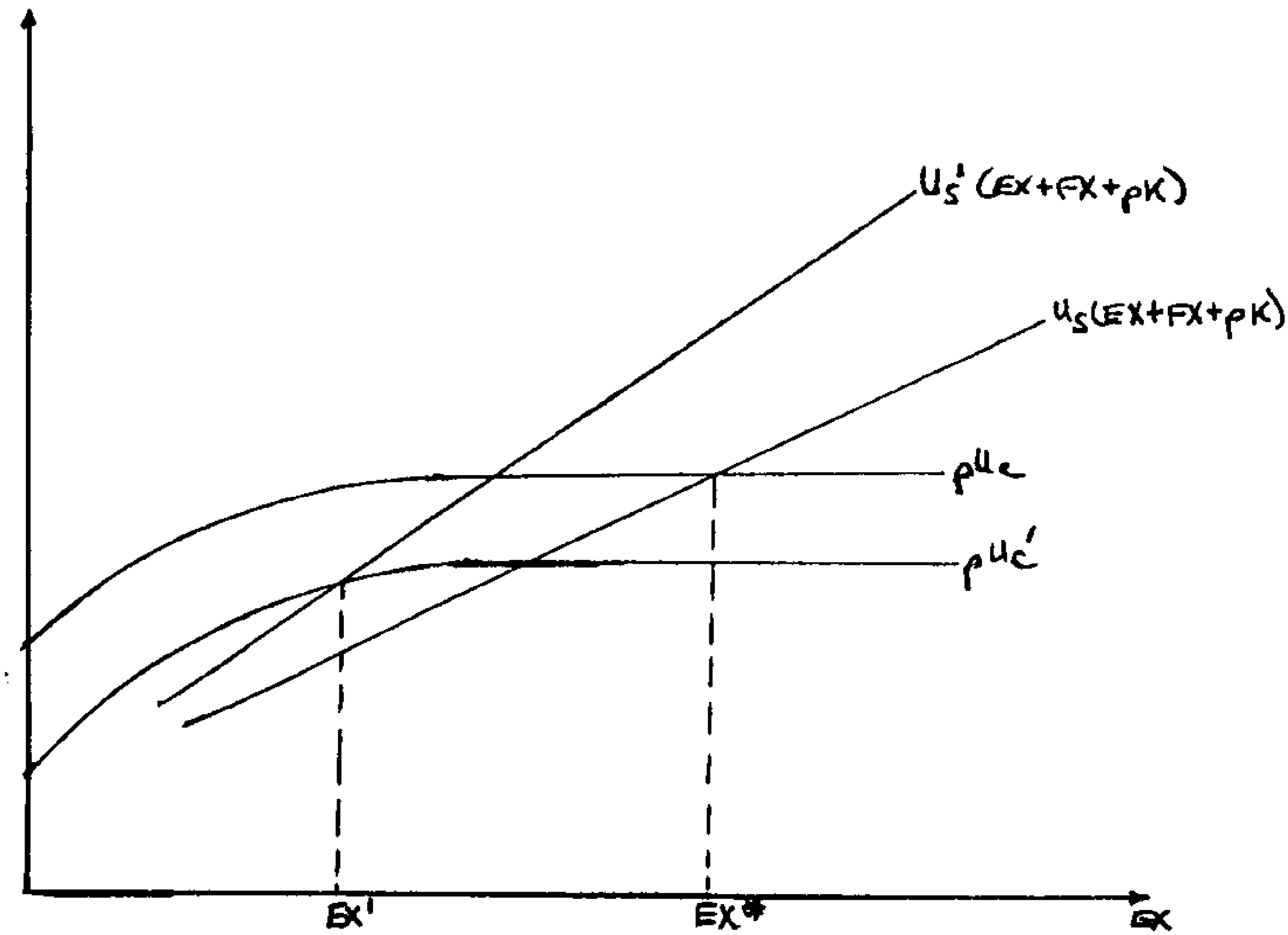


Figure 8 Economic Demonstration of Wantedness

n^*+1st pregnancy, $U_s(EX+FX+pk) > pU_c$ at EX^*. This is because pU_c is defined for $c = c^*$, then at some higher level of $c = c'$, $U_c' < U_c^*$ because of declining marginal utility of children thus the pU_c' lies below pU_c^*. On the other hand even if expenditure on the "unwanted" pregnancy were zero, income available to spend on S would decline by $(FX+p(RE,O)k)$ so that U_s should rise. Taken together these shifts imply that $U_s'(EX+FX+pk)$ will shift upward to the left and that the new intersection point EX' will be below EX^*. Thus we seem to arrive at the not unexpected result that families will spend less on unwanted children and that this new expenditure level is an attempt to reach a kind of "second best" solution given that an additional pregnancy has occurred.

One problem with this solution is that it does violence to our assumption that all children are treated equally, a basic assumption of our original model. We have arbitrarily divided pregnancies into two groups the wanted and unwanted, assuming equal expenditure levels for all wanted children but lower expenditures for those that are unwanted. Examination of Figure 7 as well as the argument behind the shifts in Figure 8 suggest that a further extension of this relaxation of the equality assumption could lead to declining expenditure per child as the number of live children in the household increased.

Consider the area p_0AG in Figure 7, this is the traditionally defined consumer surplus accruing to the consumer when p_0 is taken as given and independent of the quantity he actually consumes. It is a well established result of demand theory that a monopolist who can vary the price as quantity changes can capture this surplus for himself by lowering price along the D curve as quantity increased.

Since the price of children can be varied by households by varying EX we should not be surprised to find that in a true sequential decision model EX^* per child fell as the number of live children in the household rose. The household would continue having pregnancies until even if they spent nothing additional on a pregnancy (e.g. EX=0) the margin loss of S would outweigh the expected gain, e.g. until

$$pU_c < U_S(FX+pk)$$

this would define the true sequential stopping point which could come at a family size in excess of c^* as previously defined. Such an approach to sequential child expenditure patterning would predict that empirically expenditures on a given pregnancy would tend to be negatively correlated with the number of live children in the household prior to the pregnancy.*

---

*Notice, however, that such a finding would also be consistent with the "wanted"/"unwanted" child dichotomy.

Chapter IV- The Pregnancy Process: An Empirical Formulation

The ideal data set to test the implications of the proceeding model of interaction between the demand for children, experienced survival rates and expenditures on pregnancy would be a form of panel data which would contain information on expenditures during pregnancy for a group of women as well as data on their reproductive histories including their experience with prior pregnancies given the level expenditures in each prior situation. With such a data set, it would be possible to estimate the effect of prior outcomes, given prior expen-expenditures, on expenditures during the index pregnancy and also to investigate the relationship of the experienced efficacy of expenditures and experienced pregnancy outcomes on the decision to have additional children.

Unfortunately such data sets are not available and one has to compromise between examining survey data that contain information on the reproductive histories of individual women but little about expenditures on an individual pregnancy such as the National Fertility Survey and cross sectional data from birth certificate records which contain prenatal care information on a cohort of pregnancies in a given year. Williams (1976) has used data from the National Fertility Survey taken in 1965 to investigate the effect of experience with prior pregnancies on the decision to continue and have an additional child. In addition, she has investigated completed family size and mortality rates as determined by prior pregnancy experience and socio-economic variables as they have been measured in the National Fertility Survey and variations in birthweight using data from the National Natality Survey (1964-66). However, the emphasis of the present paper has been twofold. Although I am interested in questions similar to those investigated by Williams, I also want to take account of the fact that in this area, a substantial number of interesting and potentially expensive policy decisions rest on questions involving the efficacy of medical care in producing favorable pregnancy outcomes. For this reason, I have decided to utilize a data set containing information on the amount of prenatal care received during a specific pregnancy and the pregnancy outcome as well as the information about pregnancy history.

The data set that I have used is a sample of birth certificates from the 1970 New York City birth cohort as recorded by the New York City Department of Health. There are several advantages to using this data source.* For one thing, the number of observations is large; although I have decided to restrict the sample to the first 6 months of 1970 for reasons which will be explained below, the sample still contains in excess of 65,000 observations. Secondly, New York City collects more comprehensive statistics about each birth than most other governmental bodies: New York City birth certificates contain information not only about the birth, but information about parents' ages, their place of birth, race, parents' education and the reproductive history of the mother. Information is also available about the amount of medical care utilized during pregnancy, about the mother's health and finally, about the status of the infant at birth. In addition, New York City annually links births and death certificates so that it has been possible for me to create a data set which contains as units of observation individual births that survived

---

*Since this is cross sectional data consisting of a single cohort of pregnancies, we may have to consider potential problems resulting from censored sample biases because we will have no observations on those families whose response behavior was not to proceed with a pregnancy that would be included in the sample.

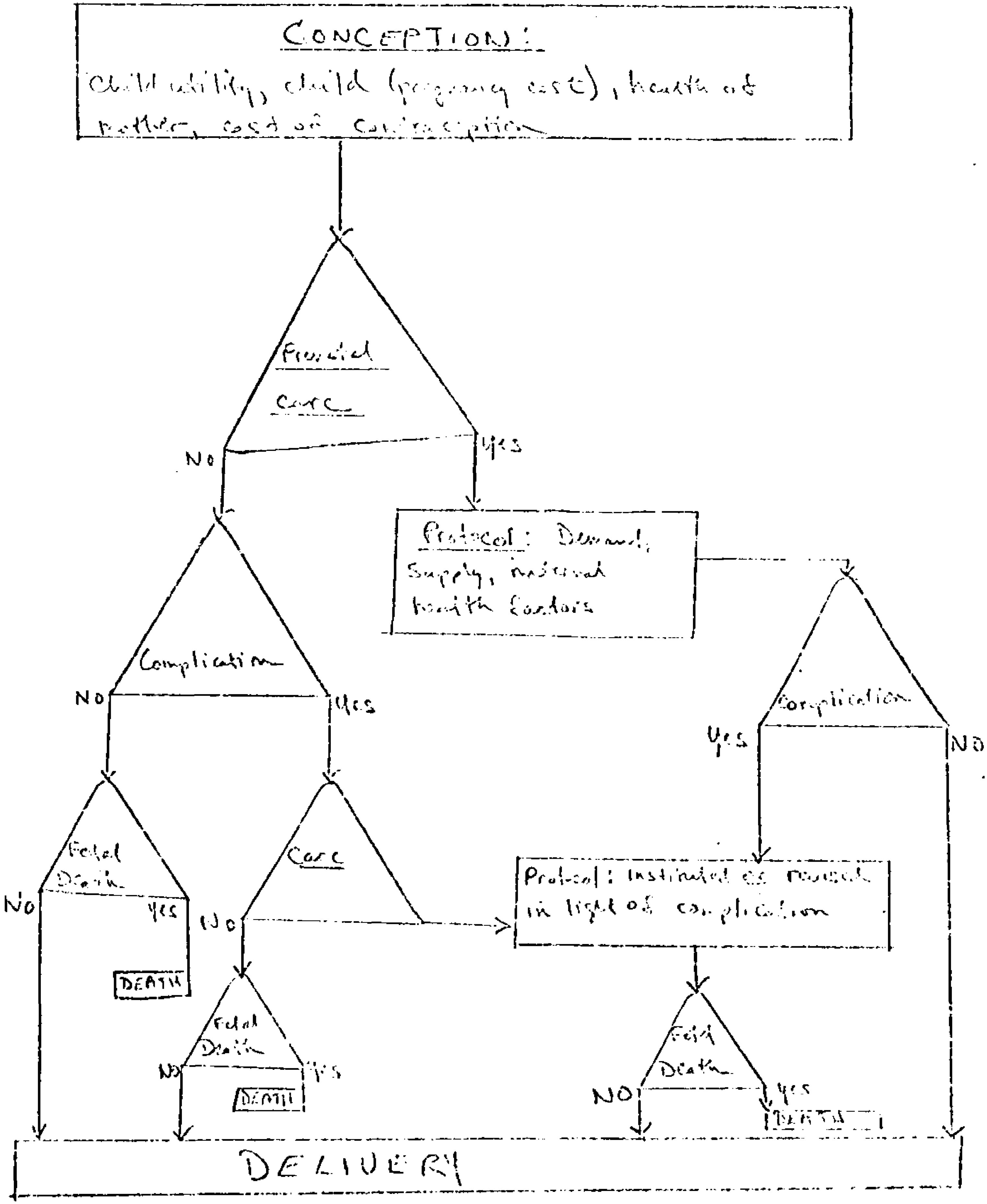
infancy as well as infant deaths with linked death certificates, giving not only the characteristics of the death, but also identifying, on a micro-basis, those circumstances of pregnancy that were associated with the death.

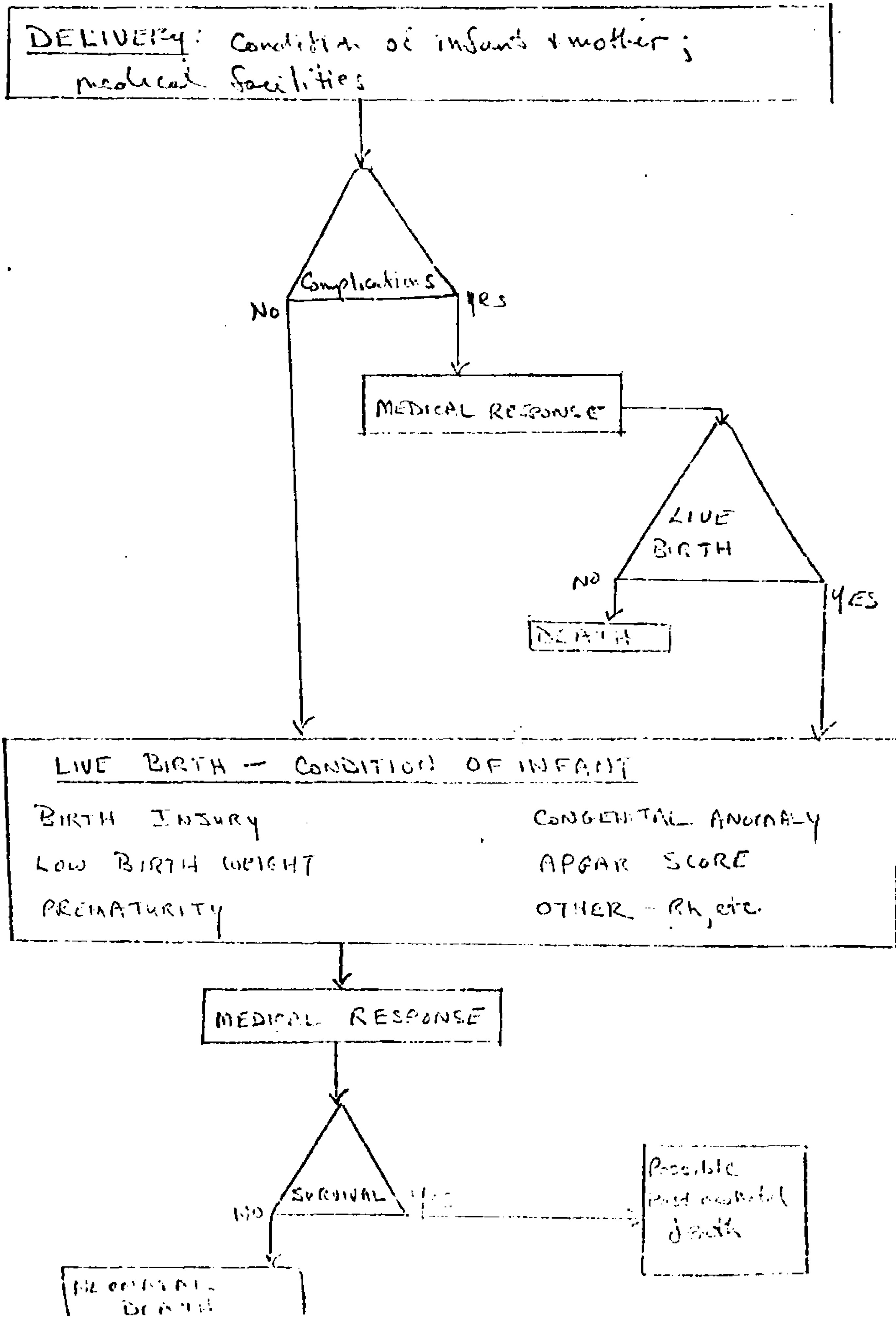
In addition to the information systematically recorded on the birth certificate, New York City has been divided geographically by the Health Department into Health Areas and Health Districts. Using information on the Health District of residence of the parents which is available on the birth certificate, I have linked the individual observations with measures of the geographic availability of facilities such as gynecologists/obstetricians, prenatal clinics in hospitals and health stations and the designation of an area for special emphasis maternal and infant/child care centers. This enables one in a fairly straightforward way to measure the effect of the availability of facilities on their utilization by individuals, which is an area of significant policy concern. Moreover, the use of individual observations tends to minimize simultaneous determination problems which might arise in such measurements because of the relationship in more aggregative units of observation between the availability of facilities and the potential demand for services.

The Pregnancy Process

Ideally, examining individual pregnancies with an eye toward measuring the interaction of events and medical care would lead to the estimation of a complex model such as the one schematized in Figure 1. This approach emphasizes the fact that pregnancy involves a production process that takes place within a given time span and that decisions

Figure 1(IV). FLOW DIAGRAM OF PREGNANCY.




DELIVERY: Condition of infant + mother; medical facilities
Complications
No
yes
MEDICAL RESPONSE
LIVE BIRTH
No
YES
DEATH
LIVE BIRTH — CONDITION OF INFANT
BIRTH INJURY
LOW BIRTH WEIGHT
PREMATURITY
CONGENITAL ANOMALY
APGAR SCORE
OTHER - Rh, etc.
MEDICAL RESPONSE
SURVIVAL
No
Yes
Possible post neonatal death
NEONATAL DEATH

made during the course of that pregnancy are made partially in response to the events that occur at various points in the time interval.

As illustrated in Figure 1, various sequences of individual behavior and medical events are possible. The individual or household, enters the pregnancy period with a set of expectations about the probable outcome, an expected demand for care and an expected level of utility resulting from the pregnancy and the decisions made during the pregnancy period. In the theoretical model, I have emphasized the relationship of these expected events to the decision to become pregnant and the "wantedness" of the child. In the flow diagram, these expectations are contained partially in the box marked conception. However, an actual pregnancy may be influenced by events which may or may not be foreseen. Such events will be reflected in subsequent behavior. For instance, prenatal care can predate a complication or be the result of a behavioral response to a complication. Similarly, prenatal care may or may not prevent complications or may or may not be able to deal with complications that arise. Lastly, even in the absence of prenatal care a successful pregnancy is entirely possible and certain complications can be weathered. In addition, it is important to note that as virtually all women are delivered by physicians in hospitals, the success of the birth at the delivery stage depends on the ability of the medical attendants and the condition of the infant at birth (which is the result of decisions made earlier in the pregnancy and presumably a large random element). It is probably only after the infant and mother are released from the hospital that the familial environm again begins to play a significant role in child health.*

Another point illustrated by the flow diagram is that attempts to study the problem by concentrating on the pregnancies which resulted in live births may lead to biased conclusions. In fact, in order to adequately measure the demand for care and a healthy child as well as the production function of a healthy child, one would have to study all pregnancies including those which end in fetal deaths somewhere along the line. This points up the importance of getting accurate information recorded about fetal deaths. Although accurate information is generally not available for early fetal deaths (miscarriages), an attempt was made to integrate observations on late fetal deaths (stillbirths) into the data base. Unfortunately, 1970 was the year when abortion was legalized in New York State. Legalization took place as of July 1st of that year. The city was not equipped to differentiate in its vital statistics for that year between elective abortions and spontaneous fetal deaths. They were recorded on the same forms and maintained on the same data tape files. Not only did this cloud the issue in distinguishing between the two events but it also may have encouraged the under-reporting of data items concerning the parents and the mother's reproductive history on the certificates

---

*It is not surprising that attempts to estimate the effect of maternal socio-economic factors on infant mortality have generally not been successful when birthweight is included as an independent variable (see Shah and Abbey [1971] as a recent example). Birthweight is an indicator of the quality of the child at birth and hence should be a function of these maternal factors. Survival of low quality infants depends critically on events during the first days of life where I would expect variations in the amount and quality of pediatric neonatal care to be the determining factors although there will be a large random element. Whether these factors should be related to such maternal factors as age, parity or even social class is not clear.

for spontaneous fetal deaths. This was because in order to assure confidentiality of abortion information little information was required or collected for abortion patients.* Moreover, abortion was only legalized up to the 20th week of pregnancy - a date many thought was arbitrary. This may have lead some physicians to under-report gestation age on some elective abortions and/or report late elective abortions as being spontaneous fetal deaths.**

The primary problem I have encountered in attempting to formulate an empirical equation system around this time-flow approach is that while the events for each pregnancy have had a definite chronology, they are frequently reported without regard to that chronology. Thus, for example, certain information concerning the health status of the mother is recorded under the general title of "Conditions Present During Pregnancy," with no indication as to whether they predated the actual conception, predated the begining of prenatal care or resulted in changes in the type of care received. On the other hand it is possible to date certain events, such as a neonatal death, because they only could have happened at a certain time.

---

*This has become an issue of some concern in the medical community. Many are concerned about the effect of repeated abortions on fertility, fecundity and even the health of the female. While much has been made of the side effects associated with oral contraceptives and IUD's, little information has been available concerning the relative risks of multiple elective abortions, particularly were it to become the dominant form of birth control.

**There has always been a strong feeling that many "elective" abortions were performed by qualified medical personnel prior to the legalization of abortion and that these were recorded as spontaneous fetal deaths with little other information. Evidence that liberalization of abortion restrictions anticipated the legal date for the easing of restrictions, can be found in the fact that the spontaneous fetal death rate for the first six months of 1970 rose significantly above the level of 1969, reversing a moderate downtrend of several years duration.

The ability to establish the actual chronology of events in a complex system such as that depicted in Figure 1 is vital in estimating the system. With only a limited number of potential exogenous variables and many endogenous variables of interest, some information on the order of events allows parts of the system to be cast in a recursive framework. In a recursive framework, events which occur earlier in a time sequence, although endogenous within the entire system, may be regarded as pre-determined in identifying and estimating relationships to explain subsequent behavior.* Thus, it is valid to include as a pre-determined variable the amount of prenatal care a mother actually received in estimating a function for infant survival since by definition all prenatal care predated the birth and therefore the potential causality runs only from prenatal care to survival. On the other hand, it is difficult to determine whether such pregnancy complications as vaginal bleeding or elevated maternal blood pressure predated the initiation of prenatal care or increased the utilization of medical care; hence, it is not possible, with the limited data available to identify equations for these different events.

---

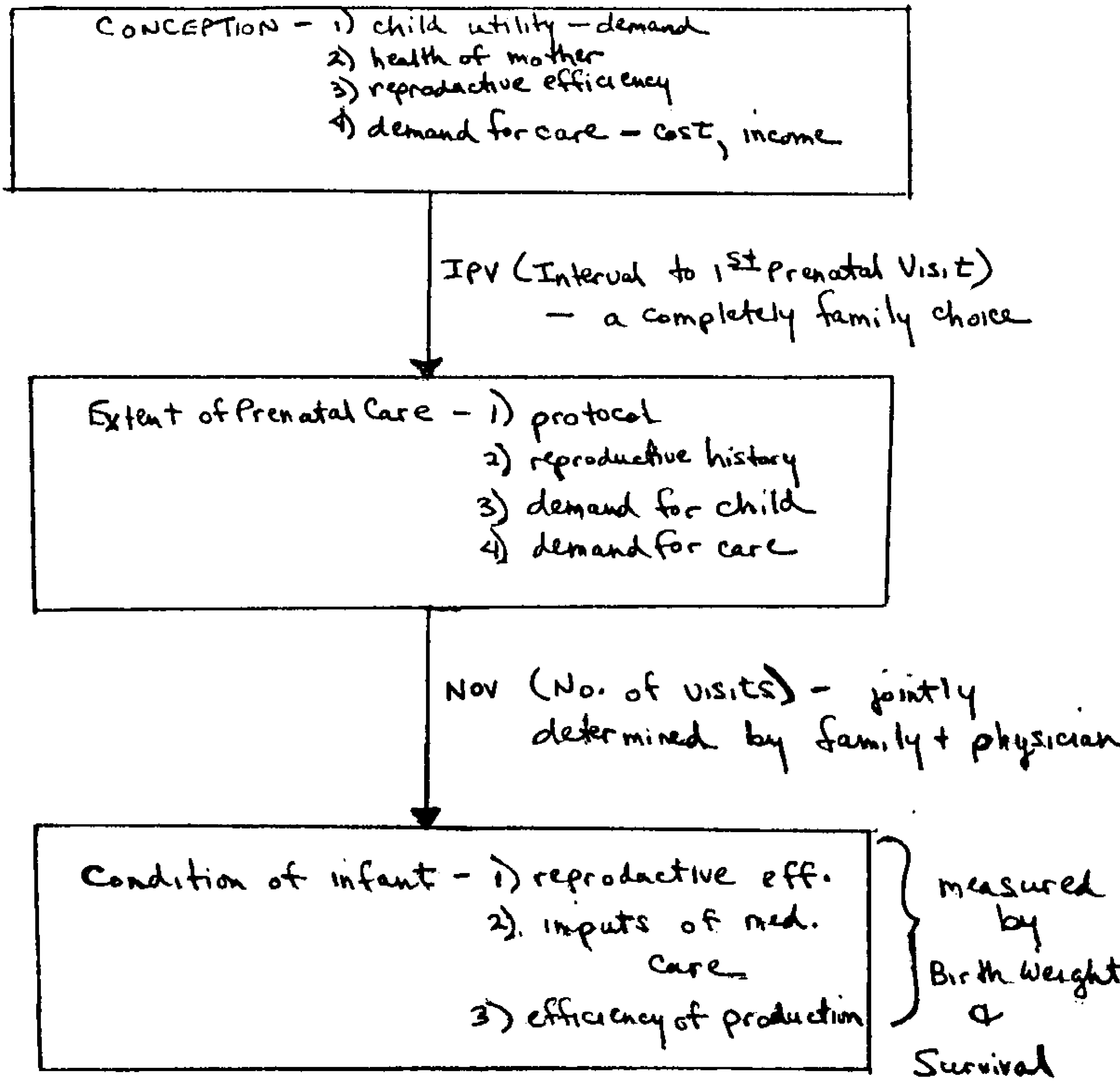
*In a fully recursive system, a sequence of endogenous variables might be represented by a causal chain without any feedback, e.g.,

$$a \rightarrow b \rightarrow c \ .$$

Knowledge of the sequence may be used to identify equations for a, b, and c as for example a will appear as a predetermined variable in the equations for b and c but neither b nor c in the equation for a. Moreover, if the disturbances in the equation for b are uncorrelated with a and the disturbances for c uncorrelated with a and b, then it can be shown that the application of ordinary least-squares estimation to the equation for each variable yields maximum-likelihood estimates (Johnston, 1972).

Because of the difficulties encountered in attempting to estimate a system corresponding to all of the possible contingencies implicit in Figure 1, a simpler system was formulated. This approach is schematized in Figure 2. It still takes into account the sequential nature of decisions and outcomes in the pregnancy process but has collapsed them into processes which can be identified and estimated with the available data. In addition, it has the added value of relating the empirical formulation much more closely to the necessarily simplified theoretical model considered earlier. Unlike the theoretical model, however, our measurement of expenditures on pregnancy is narrowed to measuring the demand for and utilization of prenatal care. This is not because other endogenous inputs such as nutrition or cigarette smoking are not considered important, but rather because as in so many other health studies information on these other inputs are not available. However, the restriction to using prenatal care as a proxy for most pregnancy expenditures (inputs) need not be viewed solely in a negative light. For one thing, prenatal care is one of the few inputs that produces health (both maternal and infant) only during the pregnancy period. Life style factors such as proper diet, adequate sanitary conditions and cigarette smoking affect individual health whether pregnant or not. However, to the extent that the prenatal care process educates mothers and encourages modifications in their life sytle favorable to health, it may have a significant effect on the pregnancy production process which operates through these other inputs. Moreover, it is important to recall that prenatal medical care and its extension to pre-pregnancy, post-partum, and family planning services have been the primary policy instruments advocated in the attempt to reduce infant mortality levels in the U.S. The Institute of Medicine Study, the establishment of Maternal and Infant Care projects, the establishment of Neonatal

Figure 2(IV). The Pregnancy/Prenatal Care Process



Intensive Care Units, and the recently voiced concern about the adequacy of existing health insurance plans to encourage adequate prenatal care (Muller, et al., 1975) all seem to indicate that the emphasis is and will probably continue to be on increasing health services utilization in order to reduce infant mortality.

The Structural Equations

The structural equations of the model, as represented in Figure 2, fall into four categories: the demand for prenatal care, the amount of prenatal care utilized, the condition of the infant at birth and survival of the infant during the first year of life. The system relies on the identifiable sequential nature of specific events to give meaning to individual relationships. Thus demand for care is measured by the interval between the last menstrual period and the first prenatal visit (INTERVAL); the amount of care by the number of visits (NUMVISIT) given the timing of the first visit; the condition of the infant at birth is measured by birth weight (WGHT) given the amount of care and timing of the first visit; and survival depends on the condition of the infant at birth, the amount and timing of prenatal care and immediate neonatal care inputs.

Demand for Care Equations

Let us now consider the demand for prenatal medical care. If medical care were a normal good, following traditional demand theory we might formulate the demand for care as

$$D_{mc} = d(P,Y,T) \qquad (1)$$

where P = price, Y = income and T is a vector of variables representing tastes. We have already noted, however, that the demand for prenatal care is a derived demand as for a factor of production and that the ultimate demand is the demand for surviving children. The demand for a factor of production would depend on its price (P), the value of its marginal product (VMP), and the desired level of output (OUTPUT).

$$D_{mc} = d(P, VMP, OUTPUT) \tag{2}$$

However, in the household production context the desired level of output depends on variables such as income and tastes as well as the shadow prices of commodities produced and hence on the parameters of the production process and the price of factor inputs.

In the context of the model developed earlier concerning the demand for endogenous expenditures on pregnancy as well as the realization that prenatal care is among the most pregnancy oriented traditional commodities consumed, we shall collapse the relationships (1) and (2) into the following

$$D_{mc} = d(P_{mc}, P_o, Y, RE, C, E)$$

where P_{mc} = price of medical care, P_o = price of other inputs, Y = income, C = current stock of live children and RE = reproductive efficiency. E represents environmental variables which directly effect the efficiency of production while not being consumed in the production process.

Let us consider the expected effect of each variable in turn. Although we have not considered price explicitly in the model, we have defined EX (expenditures on pregnancy) and S (the composite other good) in such a way that changes in EX effect the residual amount of S available to satisfy other wants. If we consider therefore a standardized unit of medical care having an associated marginal product, the higher the price (the more EX per unit) the less will be the effective amount of p (probability of survival) purchasable for a given EX and hence the lower

$\frac{\partial p}{\partial EX}$, as we have seen previously this will lead to a reduction in the amount spent on p. Moreover, if P_{mc} increases relative to the price of other standardized inputs purchasable under the umbrella commodity EX, we should not be surprised to find the effective reduction in medical care (mc) magnified. Similarly, we should not be surprised to find the amount of prenatal medical care (mc) demanded negatively correlated with the price of the other possible inputs in the production of a sucessful pregnancy.*

In Chapter 3, I have argued that Y effects the demand for inputs during pregnancy primarily through its effect on the fixed cost associated with pregnancy. Y will tend to increase the demand for inputs. Whether or not Y will be correlated with the demand for medical care depends not only on this result but the relative time intensity associated with utilization of medical care and other inputs. However, presumably one of the purposes of prenatal care which I have not directly touched on in the theoretical model is to maintain the health of the mother during and after pregnancy. In the sense in which the demand for maternal health is related to the value of time lost due to illness (Grossman, 1972) and this in turn is a function of Y, we would expect an additional positive income effect.

---

*The ease of factor substitution in production is defined by the so-called elasticity of substitution. The elasticity would measure the extent to which a firm moves along a given isoquant as relative factor prices change. The pregnancy situation we are investigating is different from the classic firm situation in 2 fundamental ways: (1) maximum output reaches a limit as $p \to 1$ and therefore marginal products may fall to zero (or even become negative) at high levels of output and (2) joint product relationships abound (e.g., good nutrition produces maternal health as well as infant health) therefore many inputs may be utilized at the level of zero marginal product in the production of p because they have positive products in the production of other commodities. In such situations, observed substitutions may reflect substitutions in the production of commodities other than infant health and responses in the infant health sector may be largely unpredictable.

Reproductive efficiency (RE) has two components which we have noted may have opposite effects on the demand for inputs generally and by extension on the demand for medical care. To the extent to which reproductive efficiency has a pure endowment effect (a pure shift in the production function), it should be negatively correlated with medical care. On the other hand, to the extent to which it is reflected in increases in the marginal products of an input, it should increase demand for those inputs whose marginal products are most augmented. In this fashion, its effect is no different than the effect of "so-called" environmental variables, except that its impact will be limited to the sphere of the production of progeny.*

Lastly, consider C. In the somewhat simplistic one-period decision model, C^* (desired family size) is determined by the cost of production and hence might not be thought to enter the function of the demand for factor input. If, however, we are operating in the sequential extension of that model and C represents live children already in the household, a different interpretation is possible. In this case D_{mc} may be interpreted not as the demand for medical care which will be the same for all pregnancies but the demand for care during a specific pregnancy. In such a formulation, the level of C has many implications. For one thing, with sequential optimization, expenditures per pregnancy may decline as the family reaches its goal stock of children - a situation similar to moving downward along a demand curve as pictured in Figure 7 of Chapter 3. Moreover, even if we aren't willing to adopt a full sequential framework with the implication of variable expenditures per pregnancy,

---

*This is not altogether true. Families who are relatively efficient producers of children may act as though this good fortune augmented their effective full income. In this way, RE may influence consumption of other goods. (For a full discussion of environmental variables in this context, see Michael (1972)).

empirically, if we are willing divide conceptions into "wanted" and "unwanted" categories then the larger is C already the more likely it is that the conception is "unwanted" in the economic sense that I have defined. Thus, the implication is that C will be negatively correlated with expenditures on medical care and presumably other inputs.*

Consider two other possible effects of the level of C on the demand for care. For one thing, C may be related to the shadow prices of time and other inputs into the production process. Hence, there may be economies of scale associated with the preparation of nourishing meals as family size increases. Older children's time may substitute for the mother's in the production of other household commodities during the pregnancy period. The presence of young children may increase the cost of prenatal visits if baby sitters are required. The expected effect of these shifts while frequently predictable for individual situations cannot usually be determined in the aggregate for most families and are not identifiable in the data set we shall use to estimate the model.

If we hold the number of pregnancies constant in a cross section, then the level of C will reflect past experience with pregnancy and possibly information about individual reproductive efficiency. As discussed previously this may have both an income and a price effect. The net effect is somewhat ambiguous. Ceteris paribus, increasing levels of C may have an income effect which encourages increases in expenditures on

---

*The problem with assuming that live birth order may be a proxy for the degree of "wantedness" has been discussed earlier (2nd footnote, pg. III-26). In that footnote, the fact that most illegitimate births occur at low birth orders and are presumably unwanted was cited as a limitation on the use of birth order as an instrument for "wantedness". While this may be true for aggregate data, in micro-data sets, such as the N.Y.C. sample used in this study, it is possible to distinguish legitimate from illegitimate births and hence the interpretation of the birth order variable presented above may have more validity.

subsequent pregnancies and a reproductive efficiency effect which discourages such increases - only if the reproductive efficiency effect is interpreted largely as representing increases in the marginal product of factors will the two effects reinforce each other.

In the data, prenatal care may be measured by number of prenatal visits (NUMVISITS) and interval from the last menstrual period to the first prenatal visit (INTERVAL). While not strictly a traditional demand measure, INTERVAL has several important attributes which make its investigation of interest. First, it is more or less accepted medical wisdom that early care is good or preferred care - that definite benefits can be derived from early monitoring of pregnancy and that specific risks can be identified early and as a consequence treated more effectively. Second, chances are good that the earlier prenatal care is initiated the more visits will be consumed. This is because there is an established protocal for scheduling prenatal visits during the course of even a normal pregnancy. Of course, individual mothers may violate this protocal, presumably without their physician's blessing, so the relationship between INTERVAL and NUMVISIT is hardly an identity.*

An aspect of this proxy demand variable that makes it of particular interest is that the interval is largely determined by the mother based on her consideration of her needs and wants. Once she enters the prenatal care process, decisions on the amount and type of care necessary will jointly be determined by the mother and her physician. The physician's interpretation of the care required by the maternal condition as well as the mother's evaluation of the information supplied by the physician

---

*One possible pattern is to seek early confirmation from a physician that the mother is indeed pregnant and then delay subsequent visits until late in the pregnancy if the mother perceives the pregnancy as being otherwise uneventful.

Table 1 (IV)

Definition of Variables

Variable Name	Definition
INTERVAL	Interval measured in days between date of mother's last menstrual period and the date of her first prenatal care visit
INTER1	Dummy variable that equals one if INTERVAL is greater than zero (i.e., if there is at least one visit)
INTER2	Dummy variable that equals one if INTERVAL is greater than 90 days
INTER3	Dummy variable that equals one if INTERVAL is greater than 180 days
INTER(350)	A transformation of the INTERVAL variable: (1) Equals (350-INTERVAL) if INTERVAL $>$ 0 (2) Equals 0 if interval equals zero
INTER(350)2	Dummy variable that equals one if INTER(350) is less than 210 days
NUMVISIT	Number of prenatal care visits
NOVISIT	Dummy variable that equals one if mother had no prenatal care visits
WGHT	Birth weight in grams
WGHTSQ	Birth weight in thousands of grams squared
AGEMOTH	Mother's age in years
AGEMSQ	Mother's age in years squared
EDMOTH	Years of formal schooling completed by mother
EDFATH	Years of formal schooling completed by father
LEGIT	Dummy variable that equals one if the child was born in wed-lock
RACEN	Dummy variable that equals one if the child is of Negro race (if either parent is Negro the child is coded as Negro)

Table 1 (IV)

Definition of Variables Cont'd.

Variable Name	Definition
FORMOTH	Dummy variable that equals one if the mother was born outside the U.S. and its possessions
PRMOTH	Dummy variable that equals one if the mother was born in Puerto Rico
CHILDLIV	Number of prevuous children born alive to mother still living at time of indexed birth
TOTLOSS	Sum of number of previous children born alive, now dead and previous fetal deaths at all gestation ages
TBO	Total pregnancy order not including current birth (i.e., TBO = CHILDLIV + TOTLOSS)
%LOSS	Per cent of previous pregnancies not surviving to date of current pregnancy (i.e. %LOSS = TOTLOSS/TBO
FIRST	Dummy variable that equals one if this is first pregnancy
LAST	Dummy variable equals one if previous pregnancy ended in a fetal death
MIC	Dummy variable equals one if mother resides in geographically defined health district which had an active Maternal and Infant Care Project health facility in 1970
CLINIC	Number of hours per week per hundred pregnancies of prenatal clinic time in all (municipal, voluntary, and MIC projects) facilities
OB/GYN	Number of obstetrician/gynecologists per hundred pregnancies (in private practice) in the health district where the mother resides
ICU	Dummy variable equals one if hospital of birth had a neonatal intensive care unit

will be important factors in determining NUMVISITS.* In fact, the existence of a protocol which the physician will attempt to encourage the mother to follow will only reinforce this inter-relationship.

The primary functions estimated for INTERVAL and NUMVISIT are

INTERVAL = f(AGEMOTH,AGEMSQ,EDMOTH,EDFATH,LEGIT,RACEN, FORMOTH,PRMOTH, CHILDLIV,TBO,%LOSS,FIRST,LAST,CLINIC,OB/GYN,MIC) (4)

and

NUMVISIT = g(AGEMOTH,AGEMS,EDMOTH,EDFATH,LEGIT,RACEN,FORMOTH,PRMOTH, CHILDLIV,TBO,%LOSS,FIRST,LAST,CLINIC,OB/GYN,MIC,INTERVAL, INTER2,INTER3) (5)

where the variables are defined in Table 1.

The rational behind the inclusion of these variables and their expected effect on demand are as follows:

(1) AGEMOTH,AGEMSQ - A nonlinear statistical relationship has been consisten observed between maternal age and reproductive loss (Nortman (1974)). If mothers are aware of the implied biological causality from age to loss, they should seek care accordingly. On the other hand, if prenatal care is significant in reducing losses and there is a nonlinear relationship between demand for care and age then the afore mentioned relationship

---

*To a large extent, particularly in the vast majority of uneventful pregnancies, prenatal care is a monitoring process. At periodic intervals, the mother's weight, blood pressure and urine are monitored primarily to check for toxemia of pregnancy, a condition potentially fatal for the mother as well as the fetus. During prenatal visits, the progress of the pregnancy, particularly fetal growth, is evaluated and the mother should be counciled on self-care at home. Nutritional guidance and usually a vitamin supplement is supplied and the mother should be alerted to the potential danger during pregnancy of smoking, indiscriminate use of drugs, and certain strenuous exercise. The patient should be alerted to monitor between visits: excessive vomiting, vaginal bleeding, persistent headaches, visual disturbance, excessive weight gain, and drainage of amniotic fluid. Since the accepted medical protocol recommends monthly prenatal visits during the first six months of pregnancy, this self-monitoring may be an important factor in the efficient utilization of prenatal medical care.

between outcomes and age will be observed.

(2) EDMOTH - Interpreted as a general environmental variable as in other health studies: as such it should be positively related to care demanded.

(3) EDFATH - Primarily serves as a proxy for permanent income of the family as such should be positively correlated with care demanded.

(4) LEGIT - If legitimate children are more desired than illegitimate child, should be positively correlated with care demanded.

(5) RACEN, FORMOTH, PRMOTH - Reflect particular differences within the population which may either be environmental, genetic or socio-economic. Of particular interest is the consistently reported finding that blacks have higher infant loss rates than whites. Once again, included in an attempt to distinguish between differences in demand which ultimately effect outcomes and biologically determined differences in reproductive efficiency.

(6) CHILDLIV, TBO, %LOSS, FIRST, LAST - A group of experience variables measuring not only the extent of experience with pregnancy (TBO), but the nature of that experience. Also takes into account stock of surviving children (CHILDLIV) as in the sequential context.

(7) CLINIC, OB/GYN, MIC - Measures of the relative availability of care. As there are no other price variables available for this sample these serve at least partially to reflect the effect of presumed differences in time and access costs on demand. Of particular interest is the effect of an MIC project on the demand for services.

(8) INTERVAL, INTER2, INTER3 - An attempt to measure and control for the effect of the protocol on the number of visits demanded. Nonlinear

terms are introduced because the protocol is nonlinear, i.e., during the first and second trimester (90-day period) of pregnancy one visit a month is recommended; in the third trimester, a visit every two weeks is recommended during the first two months and a weekly visit thereafter until delivery.

The inclusion of the INTERVAL variables in the NUMVISIT function takes account of the sequential nature of decisions during the pregnancy period. Because of the specific sequential nature of this process one can ask the question What is the nature of the determination of the amount of prenatal care demanded as a result of the patient-physician interaction given the patient's initial decision on when to seek care? Moreover, because of the presumed sequential causality running from INTERVAL to NUMVISITS, this part of the system may be regarded as recursive and therefore OLS estimation techniques may be employed.*

Outcome Measures

Two outcome measures are investigated as resulting from the pregnancy production process, birth weight and infant survival. Although most of the previous discussion has been couched in terms of infant survival as the desired outcome, the weight of the infant at birth has special attraction as an outcome measure. Birth weight is the most important single indicator of the condition of the infant at birth. Birth weight is a strong predictor of infant survival. So strong that in many studies, when birth weight is included as an independent variable, no other significant predictors of infant death have been determined (Shah and Abbey, 1972); Kessner, 1972).

---

*See footnote page IV-7.

Birth weight captures two aspects of the pregnancy process in one measure. For one thing, it is an index of maturity. During the pregnancy period, the fetus grows and develops from a single fertilized egg into a complex organism designed to cope independently with an often hostile environment. This development and maturation takes time - the gestation period. Infants born too early will not have developed the necessary body systems to deal effectively with their new environment and may thus be unable to survive or only able to survive with their impaired autonomous functions artificially supported. During the period of gestation the infant is also growing in size, gaining weight. There is a strong, stable relationship between the length of gestation period, gestation age, and birth weight. So-called intrauterine growth curves have been developed to represent the functional relationship between birth weight and gestation age (Williams, 1974).

Recent evidence has accrued to indicate, however, that birth weight is more than just a surrogate for gestation age. Using a large sample of California births (1.5 million), Williams has demonstrated that when either birth weight or gestation age is included in a sophisticated logit function to estimate their adjusted effects on survival, birth weight is by far the stronger variable, explaining 93 per cent of the variation compared with 54 per cent for gestation age. Infants who are born prematurely, by gestation age criteria, but who are of adequate or above average birth weight are much more likely to survive than infants born at term but at low birth weights.

Another reason for preferring birth weight over gestation age as an index of infant health at birth is that it is much more accurately measured. Birth weight is determined simply by weighing the infant at birth. Gestation age is typically determined (in both my New York City sample and Williams' California sample) by the time interval between the date of the mother's last menstrual period and the date of delivery. Clearly, if the data of the last menstrual period is only recorded at birth or at a prenatal visit near the birth date, there is considerable room for recall error. Only the month may be recalled or a recording error may enter the reported data if the calendar year has changed between conception and birth.*

Another problem in using gestation age is that a substantial proportion of conceptions occur prior to marriage. Since mothers are frequently reluctant to report this fact, they will tend to report the date for their last menstrual period as occurring after their marriage. This problem is so pervasive that it has become standard procedure in calculating intrauterine growth curves to disregard births at high weights for very short gestation periods as resulting from this sort of reporting error. At the other end of the gestation age spectrum are the very high ages that are reported for women who presumably missed several menstrual periods prior to actually conceiving. Although much smaller in number than the number of births at low gestation age, they do present a problem in interpretation.

Although prematurity whether measured by shortened gestation ages or low birth weight are both associated with substantially increased risk of infant mortality, follow-up studies have also indicated an

---

*See the discussion of the data editing techniques employed for an indication of the substantial errors of several years or decades encountered in using calculated intervals in the New York sample.

association between prematurity and reduced physical development and intellectual achievement at older ages of childhood. Recent research has pointed out a particular association between being a so-called "low birth weight for gestation age" infant and having learning deficiencies measured many years later in elementary and even junior high schools. These findings have lead many researchers to speak of such infants' development as being seriously compromised by prenatal "undernutrition", "malnutrition" or "fetal deprivation". However, there does not seem to be any conclusive evidence that increasing the nutritional intake of mothers of such infants during pregnancy would significantly correct this situation.*

Although there have been many studies of medical and socio-economic correlates of birth weight, none has used a production format such as I have employed in this paper. Hence, it is difficult to specify a traditional production function. This is particularly true as fetal growth in utero is a biological process which cannot be observed and

---

*Note, it is possible that many mothers of such infants are unable physiologically to deliver adequate nutrients to their fetuses in utero regardless of the maternal intake of nutrients. Such mothers might be regarded as having a reduced level of reproductive efficiency which may act to diminish the marginal product associated with the pregnancy input food.

It is conceivable that birth weight may be partially genetically determined. It has long been noted that black intrauterine growth curves are lower than those for whites. While one might be tempted to associate this finding with environmental deprivation during pregnancy, there is also substantial evidence that black infants at low birth weight have substantially lower infant mortality rates than comparable white infants. Hence, one is tempted to speculate that part of the apparent difference in intrauterine growth between the races may be genetically determined and that black infants of moderately low birth weight may be more "mature" than whites of similar birth weight.

except for animal experiments, it is socially unexceptable to experiment with outcomes by varying inputs except by specific supplementation.

It appears to be possible to only identify three endogenously determined inputs: prenatal care, maternal nutrition and several other maternal habits during pregnancy. Of these other maternal habits, smoking during pregnancy has been identified as the habit resulting in reduced birth weight which is most widely practiced. Although there is some evidence that certain drugs may have a deleterious effect on pregnancy outcome, no drug in general use has specifically been associated with low birth weight.

The mechanism by which prenatal care may affect birth weight is not very clear. Part of the prenatal care process should be to provide the mother with information on the hazards of smoking and the value of proper nutrition - not infrequently a supplemental prenatal vitamin pill may be prescribed. To the extent to which this information has value and the mother follows instructions, this aspect of the prenatal care process should beneficially affect birth weight. A second mechanism by which prenatal care may increase birth weight would be by increasing gestation age. Specific therapeutic procedures may be effective in certain cases in prolonging the gestation period and thus give the fetus more time to mature and grow. It is important to realize that routine prenatal care is primarily a monitoring procedure. Normal as well as high risk abnormal pregnancies may be monitored. Moreover, as Chard (1974) notes,

> "The value of the screening procedure is totally dependent on the efficacy of the treatment to which it leads. The literature on obstetric management is notably bereft of controlled trials, and it is almost impossible to guarantee the scientific validity of any of the well-known measures...Concrete evidence for the value of many aspects of routine antenatal care is, surprisingly, lacking..."

However, despite the validity of Chard's general observation, there are several very good reasons for including measures of the extent of prenatal care in a production function for birthweight. First, we have the data to measure the aggregate effects of prenatal care on birthweight in a large sample and perhaps shed some light on the value of such care as typically practiced. Second, this measurement is important as prenatal and perinatal care are two areas of major policy concentration on this area. And lastly, prenatal care is the only measure of endogenously determined pregnancy specific inputs present in the data set available for this study and in most other large data sets. As input measures, prenatal care is quantifiable by both the number of visits (NUMVISIT) and the interval to the first visit (INTERVAL). Both measure aspects of care that have been stressed in the medical literature as being important components of healthy pregnancies.

Unfortunately, neither measures of maternal nutrition nor health related habits such as smoking, are reported in the data. Several proxies are available, primarily the education of both parents. EDFATH may be viewed as a proxy for permanent family income. It seems reasonable to expect that if nutrition is a normal good, it should be positively correlated with income. Two aspects of nutrition are important and both should be correlated with income: the nutrition of the mother during pregnancy and the nutritional state of the mother's body immediately prior to pregnancy. EDMOTHER may serve as an index of the mother's health and nutritional state prior to pregnancy as individual health and education levels have been shown to be consistently correlated. More educated mothers may be less likely to smoke if education increases their awareness of the hazards of smoking and enables them to better translate such information into behavior favorable to health. Education may also increase the productivity with which a given budget is used to produce adequate nutrition and the productivity of the prenatal care input itself.

There are a number of variables which may reflect either exogenous biological or genetic factors, exogenous environmental factors or endogenous behavioral factors. Such variables include maternal age (AGEMOTH, AGEMSQ, TEEN), race of the child (primarily RACEN)*, and nativity of mother (FORMOTH, PRMOTH). In the case of maternal age, females at either extreme end of the fecund age spectrum may be biologically less able to produce adequately developed infants. On the other hand, there may be different environmental and behavioral factors operant in pregnancies at different ages which may impact on birth weight.

As previously discussed, race (particularly the black/white dichotomy) has been demonstrated to be associated with differences in birth weight.* However, this finding is open to any or a combination of biological, environmental or behavioral interpretations. Although neither Puerto Rican or other immigrant groups represent distinct genetic or ethnic groups, they may represent different genetic types than the average native born white mother. On the other hand, immigrant status may be associated with socio-economic factors, such as a different level of income for a measured

---

*Race referred to in the empirical work is the race of the child as reported on the birth certificate. The race of the child is derived from information given for the race of the parents. Children are coded as "white" (includes Hispanics) if both parents are white or parental race is unknown for both parents. If either parent is non-white, the child is coded to the race of the non-white parent. If either parent is negro, the child is coded as negro. For other non-white races, the race of the child is coded to the race of the father, if known.

levels of education, language barriers which make the optimal utilization of medical care difficult or different attitudes towards pregnancy and child bearing.

Several of the groups of experience variables (CHILDLIV, TOTLOSS, FIRST) have been shown to be predictors of birth weight in subsequent pregnancies (Williams, 1976). To the extent to which these experience variables measure reproductive efficiency this is to be expected and their inclusion in the current study justified. Moreover, to the extent to which they may represent a pattern of repeated behavior deletérious to fetal development they should be included in an exhaustive list of possible independent variables to determine the extent to which their value as predictors is mitigated or enhanced by subsequent behavior.

Two other variables, which I have included in the estimated function for birthweight, are legitimacy (LEGIT) and a dummy to indicate whether the mother resided in a designated Maternal and Child Care project area (MIC). Although it is difficult to make a strong case for including legitimacy as an input into a production process, one could argue that to the extent to which marriage is productive of particular efficiencies accruing to the household that some net benefit may apply in the case of pregnancy also. On the other hand, one could argue that for illegitimate pregnancies, father's education is a poor proxy for income and that as a first step in adjusting for this difference, a dummy variable to measure legitimacy should be included in any production specification. Overall, it would appear that so little is known about the pregnancy production process that the exclusion of a potentially important variable such as legitimacy a priori is unwarranted.

The MIC variable is included as an approximate proxy of the quality of care delivered. MIC projects have been established in poorer areas of the City with the goal of increasing the health of otherwise high-risk mothers and children. Many of these programs are innovative and supply not only traditional, narrowly defined medical care but also family planning and other social and educational services. To the extent to which such programs are adequately funded and truely productive, one would expect them to have a positive effect on birth weight amongst mothers whose potential risk status is being theoretically held constant in a multiple regression estimation.

Lastly, I have included gestation age in the birth weight equation. In a more fully developed system gestation age would be regarded as only an intermediate endogenous variable to be determined simultaneously with birth weight, however, in the system I am estimating I have found it difficult to identify separate functions for both birth weight and gestation age and have decided, for reasons enumerated above, to investigate birth weight. Gestation age is an important predictor of birth weight but not necessarily a productive element per se, however, the inclusion of gestation age in the birth weight equation may serve a useful purpose by allowing us to answer the question, what is the true contribution of a particular factor when gestation age is held constant? The "adjustment" value of gestation age is particularly important in accurately estimating the effect of prenatal care on birth weight. Because of the protocol implicit in the prenatal care process most women, once they begin care, will have regularly scheduled visits until they deliver. If we include a measure

If we disregard the possible error introduced by using only univariate tabulations, such variables may serve as important predictors of infant survival, maternal reproductive efficiency and even the demand for pregnancy specific inputs. At the other end of the methodological extreme the authors report in a single page on a multiple regression equation they have estimated including birth weight, as well as several totally predetermined variables as independent variables. They report that the regression equations have little use because by R^2 criteria birth weight accounted for almost all the explained variation in deaths. Clearly this interpretation and their rejection of a multiple regression technique confuses the value of the immediate predictability of death which is best provided by birth weight with an examination of the underlying biological, environmental, and behavioral causes of infant loss.*

In performing our regression experiments to predict neonatal mortality, we shall focus on three separate production formulations. An equation including only predetermined variables; one also including measures of the endogenously determined prenatal care measures; and a final formulation also including birth weight and gestation age as measures of the condition of the infant at birth. We shall be interested in determining the value of care holding the predetermined variables constant and also the value of care even after the condition of the infant is known at birth. Notice that careful interpretation of these latter

---

*An analogous experimental situation would be to perform the following two analyses of the effects of smoking on mortality. For a cohort of men, aged 55-65, compare the death rates over a five year period of those who have a history of smoking and those who have been non-smokers. The smokers will have a statistically significant higher expected mortality rate. Now, in a multiple regression, include dummy variables for both a history of smoking and a diagnosis of lung cancer at the beginning of the period - it is probable that since lung cancer has a five year survival rate near zero that the contribution of the smoking variable as a predictor of death will be near zero.

It has also been traditional when investigating infant mortality to investigate age specific rates within the infant category. The most common distinctions are made between neonatal and postneonatal mortality. The neonatal period is defined as the first 28 days of life; the postneonatal period as the period from 28 days to 1 year. One reason for making this distinction is that different causes of death may be identified with each period. Deaths caused by "certain diseases of early infancy" account for 82% of all neonatal deaths while only accounting for 4% of postneonatal deaths.* Important causes of postneonatal deaths include influenza, pneumonia, congenital malformations and accidents, accounting for over 58% of all deaths during this period. Although these differences in the causes of deaths have been the basis for the traditional distinction between neonatal and postneonatal mortality, for the purposes of this study, the age specific categories are useful because they represent deaths which take place in different settings and presumably reflect slightly different production conditions.

For the most part neonatal mortality occurs in hospitals during the period just after the infants birth. For the most part, deaths during this period are due to the production during the pregnancy of infants inadequately equiped to survive in the postnatal environment. Until fairly recently, little could be done to salvage many of these infants, now a field of medicine, neonatology has evolved to more effectively treat high risk infants. So-called neonatal intensive care units have developed to deal with high risk infants and in a city like New York an elaborate transfer

---

*The category "certain diseases of early infancy" includes birth injuries, postnatal asphyxia and atelectasis, neonatal disorders arising from certain diseases of the mother during pregnancy, and immaturity unqualified as well as other less prominent disease classifications.

network exists to move high risk neonates into hospitals with such facilities. The cost of running such units is relatively high, yet they have proliferated relatively rapidly, therefore, we will be interested in attempting to measure the effect of such facilities on infant death or survival.

As in the case of estimating production functions for birth weight, we are hampered by the lack of a well defined and measurable set of inputs into the survival production process. In fact, several production processes can be identified. One which traces the effect of only totally predetermined variables such as mother's education, age and/or prior experience with pregnancy on the death/survival outcome. One which includes endogenously determined inputs such as the amount and timing of prenatal care in the production function. And a final formulation which includes such intermediate measures of pregnancy outcome such as birthweight and gestation age. This latter relationship would provide information as to the effect of exogenous factors and endogenously determined inputs on the probability of death given the condition of the infant at birth.

One of the major failures of most previous studies of the causes of infant mortality has resulted from the failure to make these important distinctions and view the pregnancy process in the context of the several equation models that I have presented. A confounding of this problem can be seen by reading different sections of the Institute of Medicine Study (1972) discussed earlier. In a chapter reviewing the epidemiology of infant mortality, the authors present a series of essentially 2 by 2 tables relating infant mortality to such largely exogenous variables as race, birth order and mother's age - all of which are significantly related to infant mortality when used singly as individual predictors.

of the date of the first visit as one measure of the prenatal care input, then to a large extent, a measure of the number of visits becomes a fairly good proxy for the duration of the pregnancy and therefore of gestation age. Thus, if we do not include gestation age as an independent variable in estimating the effect of the amount of prenatal visits on birth weight we will substantially over-estimate the productive value of such visits as there is, for most women, a biologically determined relationship between gestation age and birth weight as discussed above.

Infant Mortality

Finally, we turn our attention to the estimation of functions for infant mortality. Although most of the previous discussion of demand has been in terms of infant survival, infant mortality as the undesirable and less likely event has been the pregnancy outcome variable traditionally studied and utilized as an index of health status and I shall not deviate from this tradition.*

---

*Of course, if we limit the possible states of being of an infant to only death or survival, then either state can be viewed as the mirror image of the other for the purposes of our study. The primary reason for the traditional concentration on infant mortality would appear to stem from a desire to identify conditions which contribute to this "undesirable" outcome and by modification of behaviour or the environment improve the chances of survival.

Note, that the concentration on infant mortality rates as opposed to survival rates has perhaps had the effect of exaggerating the apparent importance of cross sectional differences, trends and random fluctuations. Thus, reports that the infant mortality rate in Sweden is less than 50% of the rate in the U.S. are more likely to stir a response than the equally valid observation that infant survival rates in Sweden are 1% higher than in the U.S. Similarly, because of the relative size of the bases, rates of change in infant death rates over time will appear much larger than rates of change in survival rates.

equations may shed some important light on the question of the ultimate value of prenatal care in assuring infant survival.*

The investigation of postneonatal mortality is interesting because it takes us away from the largely unseeable biological production process and out of the hospital/medical environment and into the home. Although the stock of health capital the infant is born with will continue to be an important predictor of postneonatal survival (as note above the importance of congenital malformations as a cause of death during this period), infant health during this period is produced primarily within the household. Education and family income should be important inputs into child health. While race, nativity of mother and legitimacy should all largely reflect environmental differences. It will be important to measure the value of the care to the infant available under the special Maternal and Child Health projects as they may ultimately affect postneonatal mortality. Moreover, the inclusion of the two variables central to most of this investigation experience with prior pregnancy and extent of prenatal input will enable us to clarify our interpretation of the effect of these variables.

Previously, we have argued that experience as measured by such variables as TOTLOSS, FIRST, TBO, LAST AND CHILDLIV are measures primarily of biological reproductive efficiency. While prior experience with raising infants as measured by FIRST or CHILDLIV may still serve as indices of efficiency per se and of the incremental value of experience

---

*A common reservation that many denegrators of intensive prenatal care have about its value is based on a third variable explanation of its apparent effect. That is, if prenatal care is found to be associated with increased infant survival and higher socio-economic class measures, then it is argued that the observed relationship between care and favorable outcomes is really due to the third variable - socio-economic status.

in increasing efficiency, variables such as LAST or TOTLOSS (which includes fetal losses) should have less effect. Moreover, CHILDLIV can be given a different interpretation. If productive inputs particularly as summarized by income are held constant, then the greater the number of live children in the household the less the mean amount of resources per child will be and this may well have a deleterious effect on postneonatal survival.* On the other hand, if education (particularly maternal) has the effect of increasing the efficiency of household production, then we would expect it to increase survival if income and live children are held constant.

Turning to prenatal care, we note that there is little reason to expect the quantity of prenatal care received to have a direct effect on postneonatal survival. During the neonatal period, it could be argued that prenatal care was of value even after birth because it may have served as a surrogate for the adequacy of resources available to attend the birth. To the extent to which continuity of care is important during the prenatal, parturient and postpartum period, adequate prenatal care may assure adequate care during the delivery itself. Moreover, prenatal care may be important in monitoring and detecting high risk pregnancies and arranging for the appropriate care of high risk infants during and immediately after birth. Since it is not clear that any such value can be claimed for prenatal care during the postneonatal period, a finding that prenatal care appeared to be a significant predictor of postneonatal survival or death would constitute an important piece of empirical

---

*For a complete discussion of the importance of this phenomenon in less developed contries, see Wray evidence for the "third variable" hypothesis. That is, to the extent that the measured amount of prenatal care demanded during pregnancy is not productive in itself but only an index of the degree of wantedness of the child, of the amount of other productive inputs such as good nutrition into the production of a healthy child or of the ability of the parents to effectively utilize the medical care system then it may continue to be a good predictor of postneonatal death or survival even though it is of little more direct productive value during the prenatal period than during the postneonatal period.

Chapter V - Problems of Estimation

Several problems were encountered in attempting to estimate both demand and production functions. For the most part these problems stemmed from difficulties in the definition of endogenous variables, inadequate scaling of variables and the estimation of functions with dichotomous dependent variables.

A problem arose in utilizing the interval to the first prenatal visit as an indicator of demand. This was because it was difficult to assign a value to the variable when there were no prenatal visits although the interval itself was otherwise a well defined continuous variable. About 3.6 per cent of the sampled birth certificates reported no prenatal care. Although it would appear natural to define the interval as zero when there are no prenatal visits, this procedure may lead to biased estimated coefficients and confuse the interpretation of the regression estimates when considered as demand equations.

The reason for this problem may be seen by recalling that early prenatal care is considered desirable behavior by the medical profession and that therefore the interval to the first visit may be regarded as an inverse measure of demand for care (i.e., the shorter the interval the greater the demand). Consider a simple two variable example as illustrated in Figure 1. If the interval (INTERVAL), in days, is plotted on the Y-axis and income (INC) is plotted on the X-axis, and if the interval declines as income rises, there will be a scatter of points in the INTERVAL/INC plane which represent different combinations of INTERVAL and INC as illustrated in Figure 1. For those observations

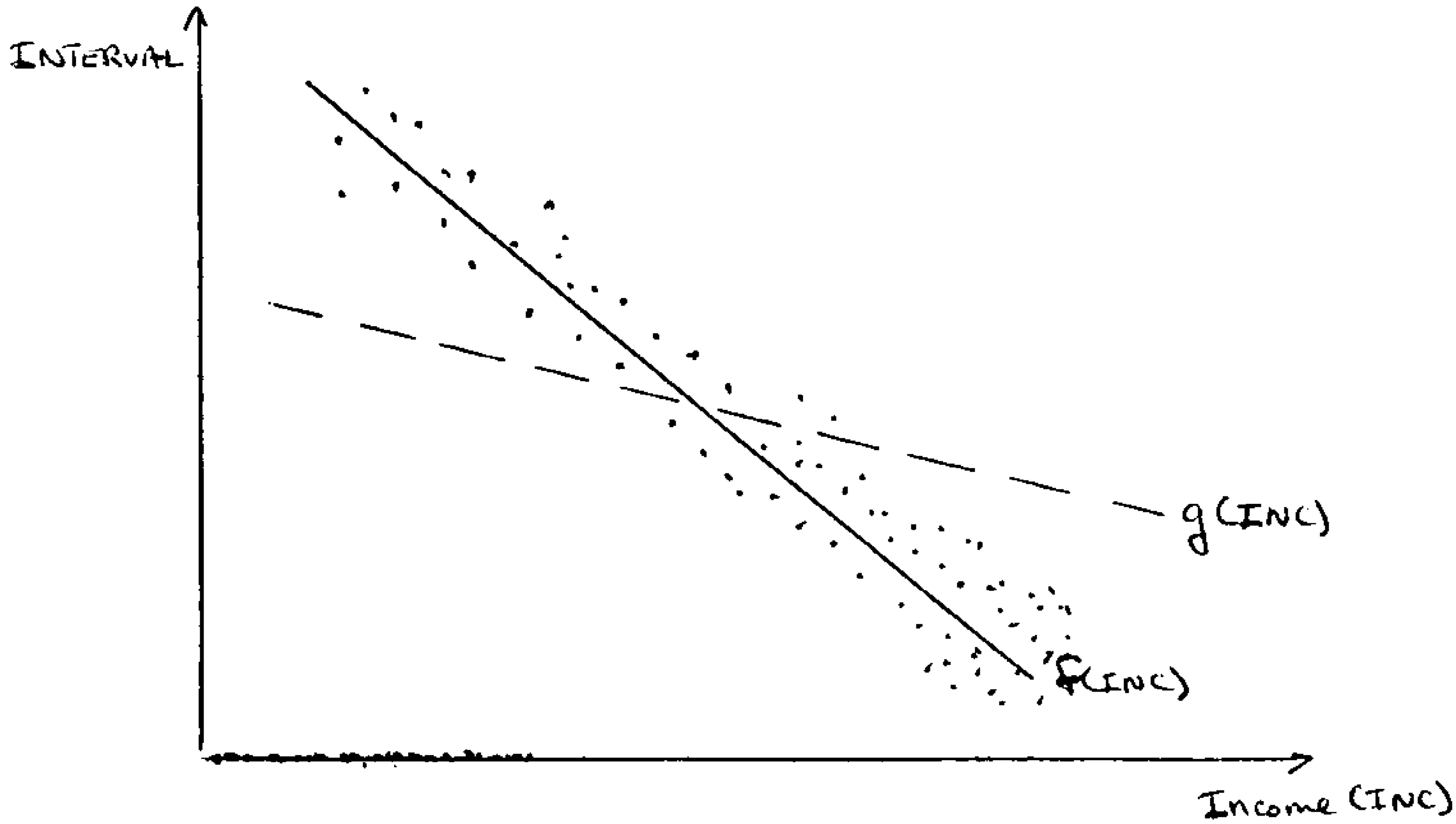


Figure 1 Possible Source of Bias in Estimation of INTERVAL Regressions

with at least one visit, one would be able to fit a function f(INC) that relates income to the size of the interval. However, if one includes observations with no visits (INTERVAL = 0) and primarily low incomes, the estimated relationship g(INC) will be much less steep and have a substantially larger standard error than the f(INC) function. In fact, based solely on g(INC) one would tend to substantially underestimate the importance of income in determining the demand for care. An alternative mathematical view of this situation is to consider the function, INTERVAL = F(INC) for all observations as being discontinuous at some low level of INC. That is

$$E(INT) = f(INC) \text{ for } INC > INC_o \quad (1a)$$

$$E(INT) = 0 \text{ for } INC < INC_o \quad (1b)$$

and moreover, E(INT) is very large for $INC \approx INC_o$ compared with E(INC) for INC much greater than INC_o.

One attempt was made to cope with this discontinuity by in fact estimating relationships that correspond to (1a) and (1b) separately. That is, for the entire sample, a function with a dependent variable corresponding to the visit/no visit dichotomy was estimated and for a sample restricted only to those pregnancies with at least one prenatal visit, a function was fit using the continuous variable INTERVAL as a dependent variable. There are, however, several limitations to this approach. One is that functions with dichotomous dependent variables present special problems of estimation, as will be discussed later. A second objection is that it may be difficult to easily interpret the results of this two step estimation procedure and that in estimating the continuous function one has essentially thrown out the information contained in the observations where there were no visits.

Two alternative strategies were employed to estimate a single function for the interval to the first visit. In one case, the interval was set equal to gestation age for all infants with no visits. The rationale behind this approach was that since nearly all deliveries occurred in hospitals and were attended by medical personnel, one could argue that the interaction between mother and physician just prior to actual delivery constituted the first (and only) prenatal visit and that gestation age was therefore, in these cases, a natural measure of INTERVAL. Notice, that the effect of this strategy is to raise those points with INTERVAL previously defined as zero off the X-axis and into the general vicinity of observations with prenatal visits. The interpretation of such an equation may be difficult because an outcome measure, gestation age, is used as a proxy for the demand for an input.*

As an example of the possible bias introduced in the estimated relationships by this approach, consider the implications of the following reasonable hypotheses: (1) both the demand for prenatal care and gestation age are functions of income and (2) therefore, most women without prenatal care will tend to be of low income and short gestation age who delivered before they ordinarily would have sought care. As a result of assigning women with no visits a shorter interval

---

*In a more complete model than is estimated in this paper, gestation age could be investigated as an intermediate outcome measure, just as birth weight is used in this paper. If, in such a model, gestation age is a function of the amount of traditionally defined prenatal care, then it would be circular to define interval to the first visit as gestation age within the context of that entire model.

than they would have planned, the coefficient of income in the estimated interval equation is biased toward zero.

An alternative strategy was to set an outer limit on the probable duration of a pregnancy and transform INTERVAL into a new variable as a function of the hypothesized potential pregnancy period. A scan of the data indicated that 350 days was a reasonable outside limit on the duration of a pregnancy (measured from the date of the last menstrual period) and the new dependent variable for visit interval was defined as:

$$\text{INTER(350)} = (350 - \text{INTERVAL}) \text{ if INTERVAL} \neq 0 \qquad (2a)$$

$$\text{INTER(350)} = 0 \text{ if INTERVAL} = 0 \qquad (2b)$$

This technique has the benefit of transforming interval into a continuous variable that will be positively correlated with the demand for care. That is, the earlier prenatal care is initiated, the larger will be INTER(350) and in the absence of any prenatal care INTER(350) will be zero. The primary draw back to using this transformation is that the size of the estimated coefficients could depend on the scaling of the variable. The constant selected, 350 days, is somewhat arbitrary and it is worth noting that a smaller constant would have moved the scatter of points with INTER(350) > 0 closer to those observations with INTER(350) = 0 while a larger constant would have spread the two groups further apart. The possible bias in the estimated coefficients that is dependent on the scaling factor appears a priori to be indeterminate.

A similar problem of definition arises when the interval to the first visit is used as an independent variable in estimating the subsequent equations in the system. In this situation, the problem associated with interpreting the variable can be seen in a relatively straight forward manner. Consider the situation where the number of visits (NUMVISIT) is a function of the interval to the first visit,

$$\text{NUMVISIT} = F(\text{INTERVAL}) \tag{3}$$

$$\text{and if NUMVISIT} \neq 0, \quad \frac{\partial \text{NUMVISIT}}{\partial \text{INTERVAL}} < 0 \tag{4}$$

For small positive intervals, the number of visits will tend to be large, but in the case of no visits both will be zero. This will tend to bias the coefficient of INTERVAL in the NUMVISIT equation towards zero.

There are several straight forward ways of dealing with this problem - two are variations of the techniques discussed above. They involve either redefining INTERVAL as INTER(350) and using that variable as the independent variable in the NUMVISIT equations or estimating a separate relationship for the dichotomout visit/no visit choice and subsequently restricting the sample for the NUMVISIT estimation only to those pregnancies with at least one visit. A final strategy which was also tried was to express INTERVAL as a series of dummy variables corresponding to the different timing of the initial visit for each pregnancy. Three such dummies have been defined INTER1 = 1 if care was initiated at least during the first trimester (90 days); INTER2 = 1 if care was instituted at least during the second trimester; and INTER3 = 1, if care was instituted in the third trimester. Thus the coefficients of these variables represent the marginal variation associated with beginning care at later dates during a pregnancy.

A problem arose in using the number of prenatal visits (NUMVISIT) as either an independent or dependent variable. This resulted from the fact that in transferring information on NUMVISIT from the actual birth certificates to the EDP files which were utilized in this study only one column was allocated to this variable. Hence, although NUMVISIT could theoretically take any value on the actual birth certificate, it only could take on the values 0-9 on the data tapes. As a substantial number of observations were coded with 9 visits (45%), this raised the possibility of introducing a bias into equations where this variable appeared.

The distribution of actual prenatal visits from a hand coded one day sample of birth registrations was available. From this sample I could determine the distribution of observations with 9 or more visits by race/nativity groupings (white, black, foreign born, Puerto Rican). The actual number of visits were tightly grouped in the range from 9-12 visits but sharply skewed towards larger numbers with a maximum at 22 prenatal visits for a single pregnancy. The mean number of visits in this range was 11.04 visits and this mean did not vary statistically among the race/nativity groupings. Accordingly, this mean, 11.04, was substituted for 9 in the actual data whenever 9 was coded on the data tape.

In order to calculate whether this limitation in the available data was important and whether the attempt to deal with it by making the described substitution was adeqiate. a dummy variable, MAXVISIT was defined. This variable equaled one when the redefined NUMVISIT equaled 11.04 and zero otherwise. When this variable was introduced into the regression estimates for birth-weight and infant death, it was only infrequently significant, had little explanatory value and did not effect the interpretation of the overall effect of the NUMVISIT dropped and will not be reported in the results sections that follow.

When NUMVISIT appeared as a dependent variable in the demand for prenatal care equations, it was similarly defined as ranging from 0 to 11.04. An alternative estimation procedure would have been to use a "Tobit" estimation technique to estimate parameters of the demand relationship. This technique, which is a maximum likelihood hybrid of probit and multiple regression analysis, is designed to provide consistent, asymptotically efficient estimates when dependent variables tend to cluster at an upper or lower bound. Using this technique, one estimates an index from which the probability of observing a non-limit value for the dependent variable and the dependent variable's expected value, conditional on a set of exogenous variables, can be determined in a single step. As the concentration of values at the limits decrease, "Tobit" estimates approach those of OLS.

Although perhaps preferred to the actual estimation technique employed, this approach was not followed because (1) there was no computer program to perform this estimation supported, debugged or available at the computer installation at which the computations were performed; (2) this estimation technique requires that the parameter values be estimated using an iterative procedure which became prohibitively expensive with a data set the size of the one available;* and (3) the results with the dummy variable MAXVISIT used as an independent variable were not sensitive to this limitation.

A similar but more siginificant estimation problem was encountered

---

*See discussion following about Logit regression estimates.

in estimating the relationships, for individual data, of infant death, neonatal death and postneonatal death. These were defined as dichotomous dependent variables equaling one if the infant died and zero otherwise. When such relationships are estimated by classical least squares, the relationship $Y = X'\beta + u$ is sometimes called the linear probability model - some will recognize it as the simplest form of linear discriminant analysis. Unfortunately, this relationship violates one of the assumptions underlying the valid application of classical least squares, namely, that the variance of the stochastic error term ($\sigma_u{}^2$) be the same for all observations (homoscedasticity). This assumption is untenable because with a binary dependent variable $E(Y_i) = P(Y_i = 1)$ and the variance of u_i is:

$$Var(u_i) = E(u_i)^2 = (X_i'\beta)(1-X_i'\beta) \qquad (5)$$

which depends on the X_i. Feldstein (1966) makes this observation in proposing such a model for perinatal mortality but notes that although classical least square estimates are inefficient in such cases they are still unbiased. He further points out following Johnston (1963) and Goldberger (1964) that the efficiency of these estimates could be improved by using an appropriate weighting scheme and generalized least squares. Feldstein argues, however, that "with a sample of nearly 17,000, the gain in efficiency does not justify the additional calculations."

Williams (1975) following Nerlove and Press (1973) rejects the linear probability model as being inappropriate in the cases of infant death. She suggests a logit specification where the logit is defined as the logarithm of the odds ratio, $\ln \frac{p(X)}{1-p(X)}$ and will range over all real values as p(X) lies between zero and one. If the logit is a linear function of some group of independent variable, $\underline{X}$ (a vector), then

$$p(\underline{X}) = \frac{e^{\underline{X}\beta}}{1+e^{\underline{X}\beta}} = \frac{1}{1+e^{-\underline{X}\beta}} \tag{6}$$

As this function is not linear in the estimated parameters ($\hat{\beta}$), it must be estimated by an iterative maximum likelihood procedure. A special program to perform such an estimation was available to the author. Unfortunately, however, because the predicted values for each observation must be calculated in each iteration to maximize the likelihood function this procedure is very expensive. Its cost increases with an increase in the number of observations, the number of independent variables and the number of iterations. Moreover, one has no clue before hand as to whether an estimation will converge within program limitations. In practice many iterations were needed for convergence and convergence frequently was not obtained.

In the program actually used (Maurer, 1973), computer core region in "K" was calculated as follows:

$$\text{Region} = \text{Constant} + \text{Buffer input/output space} + \text{Array Size} \tag{7}$$

$$\text{where, Array Size} = \frac{4 \times \text{No. of Variables} \times \text{No. of observations}}{1{,}000} \tag{8}$$

Since we were interested in relationships with approximately 15 independent variables, the estimation of these relationships for our sample of over 50,000 observations would have required over 3,500 K. Only 1,000 K would ever be available on the big 370 system I used and only jobs under 500 K could be relied on to run consistently due to the queing algorithm used to assign priorities. Accordingly, a sub-sample of observations were obtained to use in these maximum likelihood logit regressions.

This sample consisted of a 50% random sample of all deaths and a 4% random sample of all surviving births - 2,656 observations altogether. The sample was weighted in inverse proportion to the sampling fractions so that the death rates in the sample were the sample as the population as a whole. Maximum likelihood logit regressions were estimated on the weighted sample and reported in Chapter 8. Classical least squares regressions, similarly weighted, were also run on the same sample and are reported and compared with the logit estimates. Finally both these groups of estimates are compared with OLS estimates of a linear probability model of infant death using the entire sample of over 52,000 cases reported in Chapter 7.

Chapter VI - Estimation of Demand Functions

In Chapter III, we examined a theoretical model of the demand for prenatal care and developed some hvpotheses about the effects of income, education. previous nregnancy experience and the number of children already in the household on the demand for care. In this chapter, I present some empirical tests of these hvpotheses and more inclusive estimates of demand equations than are simplv imnlied bv the theoretical model. The data used, as indicated in Chapter IV, is primarilv birth certificate information from the New York Citv birth cohort for the period Januarv to June 1970. The edited data set contains 54,280 observations, approx-imately 13,000 observations having been eliminated due to insufficient recording of a sufficient number of crucial data items to justify the observation's inclusion.

Interval to the First Prenatal Visit

First, we examine the decision as to whether or not to seek prenatal care at all during a pregnancv. In Table 1 are reported the results of classical least squares regressions of the dichotomous dependent variable NOVISIT on an increasingly inclusive group of independent variables.* The coefficients in the table may be interpreted as the change in the probability that the mother had no visits per unit of the independent variable or if a characteristic is represented by a dummy variable, as the change in the probability associated with the presence of the characteristic.

Thus, we note that maternal age has a non-linear effect on the

*NOVISIT is a dummy variable that equals one if the mother had no prenatal care visits and zero if she had at least one visit.

Table 1 (VI)

NOVISIT: Dichotomous Dependent Variable that Equals One if Mother Had No Prenatal Visits

Coefficient of OLS Regression Estimates: Entire Sample
N = 54,280

(t-statistic in parenthesis)

Variables in the Equation	(a)	(b)	(c)
AGEMOTH	-0.0009 (6.04)	-0.0068 (6.34)	-0.0068 (6.28)
EDMOTH	-0.0011 (4.23)**	-.0007 (2.33)**	-0.0006 (2.20)*
EDFATH	-0.0021 (5.83)**	-0.0014 (3.98)**	-0.0006 (4.01)**
LEGIT	-0.0436 (20.97)**	-0.0403 (19.04)**	-0.0399 (18.69)**
RACEN	0.0114 (6.12)**	0.0077 (4.07)**	0.0052 (2.57)*
FORMOTH	-0.0029 (1.50)	0.0009 (0.45)	0.0007 (0.34)
PRMOTH	0.0211 (9.09)**	0.0191 (8.19)**	0.0176 (7.29)**
AGEMSQ		0.0001 (4.52)	0.0001 (4.50)
TBO		-0.0028 (1.47)	-0.0028 (1.44)
CHILDLIV		0.0108 (4.78)**	0.0107 (4.71)**
FIRST		-.0019 (0.90)	-0.0020 (0.92)
%LOSS		0.0003 (0.05)	0.0078 (1.14)
LAST			-0.0097 (2.25)*
MIC			0.0092 (4.62)**

Table 1 (VI) Cont'd.

Variables in the Equation	(a)	(b)	(c)
CLINIC			-1.0001 (3.99)**
OB/GYN			0.0006 (0.66)
$\bar{R}^2$	0.0239	0.0275	0.0281
C	0.1190	0.1876	0.1872

*Indicates coefficient significant at 5% level.

**Indicates coefficient significant at 1% level.

probability of seeking prenatal care. The probability of seeking care increases, although at a falling rate with increasing maternal age, reaching a peak at approximately 38.9 years and falling afterward. This is somewhat ironic in that most studies indicate that very young mothers are potentially at high risk of an adverse pregnancy outcome. However, this finding may also indicate that to the extent that care produces good outcomes the apparent high risk of very young mothers is due to their lower utilization of care. On the other hand, the evidence indicates that older mothers (over 35), another high risk group, are among the most likely to seek care - perhaps because of their greater experience or the relatively well publicized risk factors associated with the more advanced child bearing years.

Both mother's and father's education tend to be positively correlated with an increased propensity to seek care. The effect of father's education is somewhat stronger and I would tend to interpret this as an income effect - viewing father's education as a proxy for permanent income. Mother's education, for various reasons, can be interpreted as an efficiency variable and as such may have both positive and negative effects on the demand for care - the net effect would appear to be modestly positive.

Not surprisingly, legitimate births are more likely to have prenatal care probably reflecting a hypothesized "wantedness" effect. The dummy variables for negro, Puerto Rican and foreign born mothers, in a sense, all stand for our ignorance - they measure differences among mothers with different racial, ethnic or nativity characteristics which we cannot otherwise explain. Both negro and Puerto Rican mothers are less likely to receive care while being foreign born has no statistically significant effect. Interestingly, the size of the coefficient on RACEN (negro mother) is reduced by over one half by the introduction of experience and availability of care variables in regression (b) and (c). This would tend to indicate that there are less "unexplained" factors affecting the utilization of prenatal care facilities by blacks than might be attributed to the racial characteristic itself.

The coefficients of the experience variables TBO, CHILDLIV, FIRST, %LOSS and LAST introduced in regressions (b) and (c) lend limited support to the hypothesized effects of experience. In particular, the probability of not seeking care increases with each live child already in the family. A finding consistent with either an experience or a wantedness effect. However, the probability of seeking care is higher for those mothers whose most recent previous pregnancy ended in a fetal death - probably an experience effect.

The variables associated with the availability of care had a somewhat surprising and mixed effect on the care/no care decision. The availability of clinic time appeared to increase the probability of seeking care, however, the availability of obstetrician/gynecologists had no significant effect. Most surprisingly, the availability of a Federally funded Maternal and Infant Health Care Project facility in a neighborhood (Health District) was associated with an increased probability of failing to seek care. One possible explanation for this finding is that there is a simultaneity problem associated with the interpretation of this result. That is, MIC projects are established primarily in areas where the utilization of prenatal care facilities is already low.

This explanation would be more convincing if these MIC projects had been established in the period immediately prior to 1970. However, the MIC program had been underway since 1962 so that adequate time had elapsed within the individual Health Districts to encourage increased use of prenatal care facilities by 1970. Alternatively, one could argue that these projects were started primarily in areas where the availability of alternative providers of care was so limited that even after the MIC facilities had been in existence for 8 years they could not, at the level at which they were operating, compensate for the lack of other facilities in a given geographic area.*

Tables 2, 3, and 4 report the estimated equations of the demand for care when various measures of the interval between the date of the mother's last menstrual period and the date of her first visit are used as dependent variables. The regressions reported in Table 2 were run only on that sub-sample of 52,552 mothers who had at least one visit. In Table 3, INTER(350) is the dependent variable and in Table 4, INTER(GEST) is the dependent variable. As the coefficients of all nine regressions reported in the three tables are remarkably similar, only overall effects will be discussed and exceptions to this general similarity noted.

As reported with regard to Table 1, the relationship between maternal age and the length of the interval to the first visit appears

---

*The simple correlation coefficients between MIC and CLINIC, and MIC and OB/GYN are .49 and -.17 respectively indicating that MIC areas did tend to have a somewhat reduced availability of private physicians but this may have been well compensated for by the presence of clinic facilities. It should be noted that neither CLINIC nor OB/GYN completely adequately measure the amount of provider capacity in a Health District.

Table 2 (VI)

INTERVAL: Interval Measured in Days between Date of Mother's Last Menstrual Period and the Date of Her First Prenatal Visit

Coefficients of OLS Regression Estimates For the Sample of Mothers Making at Least One Prenatal Visit, N = 52,552

(t-statistic in the parenthesis)

Variables in the Equation	(a)	(b)	(c)
AGEMOTH	-0.9389 (20.86)**	-9.0350 (26.61)**	-8.9093 (26.25)**
EDMOTH	-1.0116 (11.46)**	-0.5745 (6.53)**	-0.6114 (6.95)**
EDFATH	-2.7976 (24.31)**	-2.2317 (19.47)**	-2.2019 (19.23)**
LEGIT	-27.9751 (41.49)**	-24.1184 (35.53)**	-23.2206 (33.98)**
RACEN	28.8665 (48.04)**	25.9259 (43.24)**	24.1485 (38.15)**
FORMOTH	10.3381 (16.98)**	13.1904 (21.75)**	12.7555 (21.01)**
PRMOTH	25.2312 (38.85)**	24.1240 (32.67)**	22.0326 (28.87)**
AGEMSQ		0.1331 (21.75)**	0.1312 (21.46)**
TBO		0.6227 (1.04)	0.5849 (.97)
CHILDLIV		4.9878 (6.95)**	4.9310 (6.87)**
FIRST		-1.0475 (1.55)	-1.1454 (1.69)
%LOSS		-8.2524 (4.39)**	-3.8456 (1.79)
LAST			-6.1099 (4.50)**
MIC			4.7311 (7.64)**

Table 2 (VI) Cont'd.

Variables in the Equation	(a)	(b)	(c)
CLINIC			0.0160 (3.09)**
$\overline{R}^2$	0.1946	0.2156	0.2176
C	198.9807	294.9729	290.9597

*Indicates coefficient significant at the 5% level.

**Indicates coefficient significant at the 1% level.

Table 3 (VI)

INT(GEST): Interval Measured in Days between Date of Mother's Last Menstrual Period and Date of her First Prenatal Visit; Equals Gestation Age if Mother had No Prenatal Visits

Coefficients of OLS Regression Estimates: Entire Sample
N = 54,280

(t-statistic in parenthesis)

Independent Variable	(a)	(b)	(c)
AGEMOTH	-1.0296 (21.72)**	-9.6025 (26.93)**	-9.4606 (26.55)**
EDMOTH	-1.1515 (12.41)**	-.6583 (7.12)**	-.6903 (7.46)**
EDFATH	-3.1115 (25.68)**	-2.4811 (20.57)**	-2.4488 (20.32)**
LEGIT	-31.9660 (45.66)**	-27.8431 (39.55)**	-26.9004 (37.97)**
RACEN	29.6066 (46.90)**	26.2541 (41.69)**	24.1167 (36.31)**
FORMOTH	9.7840 (15.15)**	13.0608 (20.31)**	12.6049 (19.58)**
PRMOTH	27.3092 (34.98)**	25.9609 (33.60)**	23.6419 (29.60)**
AGEMSQ		.1399 (21.74)	.1378 (21.43)
TBO		.4286 (.68)	.3900 (.62)
CHILDLIV		5.8904 (7.84)**	5.8151 (7.75)**
FIRST		-1.3563 (1.90)	-1.4409 (2.202)*
%LOSS		-8.5327 (4.30)**	-3.3792 (1.49)
LAST			-7.0637 (4.92)**

Table 3 (VI) Cont'd.

Independent Variables	(a)	(b)	(c)
MIC			-7.0637
			(4.92)**
CLINIC			5.9563
			(9.16)**
OB/GYN			.0768
			(.24)
Constant	213.15	314.42	310.15
R^2	.202	.225	.226

*Indicates coefficient significant at 5% level.

**Indicates coefficient significant at 1% level.

Table 4 (VI)

INTER(350): 350 - Interval Measured in Days between Date of Mother's Last Menstrual Period and the Date of her First Prenatal Visit; Equals Zero if Mother had No Prenatal Visits

Coefficients of OLS Regression Estimates: Entire Sample
N = 54,280

(t-statistic in parenthesis)

Independent Variable	(a)	(b)	(c)
AGEMOTH	1.0979 (20.99)**	10.1455 (25.78)**	10.0024 (25.42)**
EDMOTH	1.2372 (12.09)**	.7012 (6.88)**	.7314 (7.16)**
EDFATH	3.2645 (24.42)**	2.5799 (19.38)**	2.5523 (19.02)**
LEGIT	35.7779 (46.32)**	31.3818 (40.39)**	30.4144 (38.89)**
RACEN	-30.7679 (44.18)**	-27.1017 (38.99)**	-24.8127 (33.80)**
FORMOTH	-9.6078 (13.48)**	-13.1955 (18.59)**	-12.7261 (17.91)**
PRMOTH	-28.9406 (33.60)**	-27.4190 (32.15)**	-25.0155 (28.37)**
AGEMSQ		-.1468 (20.67)**	-.1447 (20.40)**
TBO		-.2356 (.34)	-.1991 (.29)
CHILDLIV		-6.7611 (8.15)**	-6.6779 (8.06)**
FIRST		1.4600 (1.86)	1.5506 (1.97)*
%LOSS		8.4228 (3.84)**	2.6557 (1.06)
LAST			7.8569 (4.96)**

Table 4 (VI) Cont'd.

Independent Variable	(a)	(b)	(c)
MIC			7.8569 (4.96)**
CLINIC			-6.5521 (8.97)**
OB/GYN			-.1000 (.28)
Constant	127.35	20.67	24.91
$\bar{R}^2$	.194	.216	.218

*Indicates coefficient significant at 5% level.

**Indicates coefficient significant at 1% level.

nonlinear. However, the interval is shortest for mothers of 34 to 35 years despite the fact that the risk of an adverse outcome rises steeply as maternal age increases beyond this age. As in the case of the visit/no visit equation, both mother's and father's education are negatively correlated with the length of the interval but the effect of father's education is much stronger, particularly after taking account of the experience variables. Thus in regressions (b) and (c) of each of Tables 2,3, and 4, the coefficient of father's education is approximately 3.6 times as great as the coefficient of mother's education. This result is consistent with a hypothesized strong income effect as captured in father's education and a weaker reproductive efficiency effect as captured in mother's education.

Illegitimate births have their first prenatal visit, if any, delayed about a month as compared with legitimate births - although the measured delay is smallest for that sub-sample of births which only include pregnancies with at least one visit. Black, Puerto Rican and other foreign born mothers also delayed first visits. The size of the coefficients of RACEN and PRMOTH are comparable to the coefficient of LEGIT, although more frustrating because they are not as amenable to explanation.

The coefficients of the experience variables indicate no specific birth order effect per se, although the coefficients reported in equations (c) of Tables 3 and 4 would intend to indicate a slight shortening of the interval associated with first births. The coefficients of CHILDLIV in all three tables tends to indicate that the presence of live children or alternatively previously successful pregnancies tend to substantially extend first visit intervals from five to seven days per live child. On the other hand, the effect of negative experience seems to be concentrated in only the outcome of the previous pregnancy. Although a history of bad experience as measured by %LOSS tends to reduce the first visit interval, the coefficient on %LOSS looses significance when LAST is introduced into the regressions. This is probably due to the high correlation between %LOSS and LAST (simple correlation coefficient equals .67) and the fact that recently experienced losses as captured by LAST may have a more substantial impact on behavior than those experienced earlier in the child bearing life cycle which may be picked up in %LOSS. Moreover, LAST measures a specific pregnancy related loss (a fetal death) while %LOSS includes child losses at any age as well as fetal losses. The coefficient of LAST is not insubstantial in any set of regressions and in fact is larger than the coefficient of CHILDLIV in all regressions in which they both occur. The significant negative coefficient on LAST is consistent with interpreting previous failures as indicating a reduction in reproductive efficiency of the production function shift type rather than representing a change in the value of the marginal product of prenatal care inputs.

Turning our attention to the availability of care variables (MIC, CLINIC, OB/GYN), we note that, as in the case of the visit/no visit regressions reported in Table 1, being located in a designated MIC project area tends to be associated with a decreased demand for care (as measured by a lengthened first visit interval). The availability of obstetrician/gynecologists has no effect on the interval and the availability of clinic time is inconsistent. The coefficient of CLINIC is only significant in the interval regressions restricted to those mothers who had at least one visit (Table 2, Regression (c)). In this case, it has a positive sign indicating that an increase in clinic hours per pregnancy in a Health District is associated, somewhat paradoxically, with an increase in the first visit interval. This finding is somewhat contrary to expectations particularly as this variable was previously found to be positively associated with initiating care in the first place (Table 1). To some extent the opposite effects of this variable as reported in Tables 1 and 2 may account for its failure to achieve significance in either Table 3 or 4. Although I have tended to view an increase in CLINIC as being associated with a decrease in the price of prenatal care (i.e., the more clinic hours are available per pregnant mother, the lower the net cost of a visit), it is at best a very imperfect measure of either time or information costs as the numerator, clinic hours, doesn't adequately capture differences in individual clinic production capacities and the denominator, pregnancies in a health district, is only a measure of potential demand for and not the actual use of a facility. The ratio of actual clinic capacity to their actual use would be a better surrogate index of the availability (cost) of clinic care to an individual pregnant mother.

Number of Prenatal Visits

In considering the demand for prenatal care as measured by the number of prenatal visits (NUMVISIT), we shall primarily be concerned with contrasts between demand functions estimated solely on predetermined variables and those which take the length of the interval to the first visit into account. The latter group of relationships will enable us to assess the impact of physicians' input on the demand for care. These two sets of regressions are presented in Table 5.

Equations (a) through (c) in Table 5 contain only variables representing the socio-economic characteristics of households, pregnancy experience and the availability of care facilities, and as such, they represent reduced form demand equations. As in the case of the interval estimates, the number of visits is a nonlinear function of mother's age reaching a maximum at about 36 years - an age just below the average woman's final years of fecundity, the period when the risk of abnormal pregnancy outcomes is greatest. Mother's and father's education both tend to increase the number of visits but the importance of EDFATH relative to EDMOTH is reduced - the elasticity at the mean of EDFATH is .11 and of EDMOTH .06 (based on regression (c)). The signs of the coefficients on LEGIT, RACEN, FORMOTH and PRMOTH are as expected and indicate a substantial amount of variation in demand attributable to these characteristics which cannot really be well explained. The experience variables all behave as expected. There appears to be a reduction in the number of visits of approximately .3 per live child in the home while first pregnancies receive an added amount of care. As in the INTERVAL regressions, %LOSS, the cumulative experience variable, while it has a positive effect when entered alone is dominated by the LAST (last pregnancy ended in fetal death = 1) when both are entered into the equation.

The coefficients of the variables representing the availability of care providers are consistent with the results of the INTERVAL regressions. Mothers in MIC designated areas have .4 fewer visits while the availability of clinic hours tends to increase the number of visits, although not

Table 5 (VI)

NUMVISIT: Number of Prenatal Care Visits

Coefficients of OLS Regression Estimates:
Entire Sample, N = 54,280

(t-statistics in parenthesis)

Independent Variables	(a)	(b)	(c)	(d)	(e)	(f)
AGEMOTH	0.271 (14.62)**	0.360 (19.07)**	0.355 (18.85)**	0.126 (8.21)**	0.067 (4.46)**	0.097 (6.23)**
EDMOTH	0.060 (12.23)**	0.041 (8.41)**	0.041 (8.28)**	0.026 (6.53)**	0.019 (4.92)**	0.022 (5.53)**
EDFATH	0.096 (15.04)**	0.073 (11.37)**	0.073 (11.33)**	0.014 (2.68)**	-0.0006 (0.11)	0.008 (1.47)
LEGIT	1.328 (35.42)**	1.255 (33.69)**	1.224 (32.65)**	0.450 (14.69)**	0.346 (11.41)**	0.381 (12.14)**
RACEN	-1.116 (33.51)**	-0.973 (29.18)**	-0.848 (24.10)**	-0.312 (10.81)**	-0.135 (4.77)**	-0.244 (8.34)**
FORMOTH	-0.121 (3.56)**	-0.268 (7.88)**	-0.254 (7.46)**	0.032 (1.14)	0.112 (4.09)**	0.063 (2.27)**
PRMOTH	-0.076 (23.71)**	-0.875 (21.39)**	-0.785 (18.58)**	-0.221 (6.40)**	-0.066 (1.94)**	-0.158 (4.51)**
AGEMSQ	-0.004 (12.73)**	-0.005 (14.63)**	-0.005 (14.49)**	0.002 (6.07)**	-0.0008 (2.80)**	-0.001 (4.40)**
CHLDLIV		-0.299 (22.88)**	-0.286 (7.19)**	-0.101 (3.15)**	-0.093 (2.95)**	-0.098 (3.01)**

Table 5 (VI) Cont'd.

Independent Variables	(a)	(b)	(c)	(d)	(e)	(f)
FIRST		0.225 (6.05)**	0.229 (6.08)**	0.189 (6.21)**	0.183 (6.11)**	0.192 (6.24)**
%LOSS		0.250 (3.64)**	0.030 (0.25)	0.008 (0.08)	-0.048 (0.50)	-0.014 (0.14)
TBO			-0.006 (0.18)	-0.015 (0.57)	0.00001 (0.0000)	-0.007 (0.24)
LAST			0.302 (3.98)	0.103 (1.67)	0.074 (1.23)	0.086 (1.39)
MIC			-0.417 (11.91)**	-0.251 (8.88)**	-0.229 (8.16)**	-0.247 (8.61)**
CLINIC			0.002 (7.52)**	0.002 (8.48)**	0.002 (9.81)**	0.002 (8.91)**
OB/GYN			-0.024 (1.39)	-0.018 (1.35)	0.021 (1.54)	0.020 (1.12)
INTER1				9.099 (141.66)**		
INTER2				-1.248 (48.24)		-0.678 (13.38)
INTER3				-2.770 (92.89)		0.282 (7.13)

Table 5 (VI) Cont'd.

Independent Variables	(a)	(b)	(c)	(d)	(e)	(f)
INTER(350)					0.030 (110.21)	
INTER(350)2					0.119 (3.21)	
INTERVAL						-0.024 (50.57)
R^2	0.133	0.149	0.152	0.447	0.460	0.347
Constant	1.690	0.657	0.748	-2.245	-0.175	8.963
N	54280	54280	54280	54280	5428	52552

*Coefficient statistically significant at 5% level.

**Coefficient statistically significant at 1% level.

[a]Regression (f) restricted to these observations for which NUMVISIT > 0.

Table 6 (VI)

Means and Standard Deviations of Dependent and Independent Variables - For the Entire Sample N=54,280

Variable	Mean	Standard Deviation
MOTHAGE	25.16	5.51
EDMOTH	10.94	3.37
EDFATH	11.90	2.60
LEGIT	0.79	0.41
RACEN	0.29	0.46
FORMOTH	0.22	0.42
PRMOTH	0.17	0.38
TBO	1.31	1.74
CHILDLIV	1.11	1.48
TLOSS	0.19	0.63
FIRST	0.41	0.49
%LOSS	0.07	0.20
AGEMSQ	663.38	300.21
MIC	0.37	.48
CLINIC	53.49	52.54
OB/GYN	.50	.83
LAST	0.06	0.23
INTER(350)	219.46	71.71
INTER1	0.97	0.18
INTER2	0.61	0.48
INTER3	0.18	0.39
INTER(350)2	0.36	0.48

Table 6 (VI) Cont'd.

Variable	Mean	Standard Deviation
NUMVISIT	7.99	3.30
novisit	0.03	0.18
WGHT	3176.14	569.27
WGHTSQ	10.41	3.44
GESTN	277.89	24.15
ICU	0.43	0.49
DEATH	0.014	0.13
NEONATAL DEATH	0.014	0.12
POSTNEONATAL DEATH	0.004	0.06

substantially (the elasticity of CLINIC in regression (c) is .01). Surprisingly, particularly considering the positive sign on CLINIC, the coefficient of OB/GYN has a negative sign although it is not significant.

Turning our attention to regressions (d) through (f) in Table 5, we examine the effect of including the interval to the first visit in the NUMVISIT estimates. In regression (d), I have used a number of incremental dummy variables to measure the effect of the length of the interval and to take account of the nonlinearity of the protocol which obstetricians are supposed to follow in scheduling expectant mothers for prenatal care.* In regression (e) I use the transformed variable INTER(350) and a dummy variable, INTER(350)2, to account for nonlinearity. In regression (f), the sample is restricted to only those mothers who had at least one prenatal visit and the definition of INTERVAL as the actual interval to the first visit is followed. INTER2 and INTER3 once again are introduced to account for nonlinearities.

The effect on the coefficients of the predetermined variables of including these various measures of the interval to the first visit is substantial. In the case of mother's age, there is a tendency for the peak at the maximum number of visits to shift to older ages. This is particularly pronounced in regression (f) where the maximum number of visits is reached at age 48 – an age at which most women are no longer

---

*Recall that the protocol is nonlinear, i.e. during the first and second trimester (90-day period) one visit a month is recommended; in the third trimester, a visit every two weeks is recommended during the first two months and a weekly visit thereafter until delivery. In terms of the marginal dummy variables INTER1, INTER2 and INTER3, we would expect the coefficient of INTER1 (equals 1 if INTERVAL is greater than zero) to be greater than either INTER2 or INTER3 but that INTER3 should be greater than INTER2 since it is primarily during the third trimester that the nonlinearity becomes most acute and the period between visits progressively shorter.

fecund. The tendency of the number of visits to continue to rise amongst older pregnant women, particularly those that have made an initial physician contact, would indicate that physicians, at least, appear to act as though encouraging higher risk mothers, as measured by age, to obtain more prenatal care.

The effect of mother's and father's education is also altered in these specifications. Father's education which was seen to be more important in determining the interval to the first visit is only significant in specification (d). Moreover, in all three specifications, (d) through (f), mother's education is significant and had a larger positive effect. This finding tends to support the notion that there is a productivity effect associated with mother's education which would tend to increase the productivity of prenatal visits and encourage their purchase. The absence of a strong father's education effect, which we have viewed as a proxy for an income effect may partially result from the way in which prenatal care and obstetrical delivery is priced by many private physicians. Most private physicians charge a single fee for prenatal care and obstetrical delivery - there is generally no individual charge for each prenatal visit. For this reason, income may effect the decision as to whether to purchase the entire package of care but not individual pieces of it as they cannot usually be purchased separately.* Pricing in clinics and MIC centers is varied and while these providers would tend to charge on a per-visit basis, most users of these facilities are

---

*Perhaps the only way to purchase these services separately would be to start care very late and persuade the physician to reduce the fee accordingly as fewer prenatal visits would be expected. If this was a widely followed practice then we might expect income proxy variables, such as father's education, to affect the interval but not the number of visits with the interval held constant.

probably substantially, if not totally, subsidized by Medicaid and other public funds and so unaffected by the method of charging for prenatal care.

The introduction of interval measures in the number of visits regressions substantially reduces the size of the previously largely unexplained differentials associated with legitimacy and the race/ethnicity/nativity dummies. In fact, the coefficient on FORMOTH becomes positive and significant in regressions (e) and (f). This indicates that foreign born mothers tend to have slightly more visits than native born mothers once one takes account of the longer interval they wait before initiating care. In general, it would appear that as much as 80% of the apparent reduction in the number of prenatal care visits associated with illegitimacy, or being a black or Puerto Rican mother can be attributed to the longer intervals these mothers wait before initiating care and not to substantially reduced levels of care once they make contact with the medical care system.

The bad experience variables, %LOSS and LAST, as well as TBO are not significant determinants of the number of visits holding the interval constant. This would tend to indicate that physicians are less likely than expectant mothers to modify their care protocol with regard to a specific pregnancy to take account of a history of previous pregnancy losses. On the other hand, both CHILDLIV and FIRST remain significant in regressions (e) and (f), although the size of their coefficients are reduced by about 50%. It is tempting to view these results as reflecting a wantedness or maternal time effect, particularly when contrasted with the lack of significance of the other experience measures. Thus, if demand is particularly high for the first child and declines with each successive birth, one would expect the pattern of the demand for care to be as reflected in the coefficients of CHILDLIV and FIRST in regressions (d) through (f). On the other hand, it is equally plausible to argue that the presence of children in the home effectively changes the value of the mother's time in such a way so as to discourage the mother from leaving home to seek prenatal care as frequently as she might do if there were no children present. Lastly, it is possible that these findings represent a "portfolio" effect (although this might better be captured in the coefficient of TBO), that is, people who plan on having large families other things equal may spend less per child, less per pregnancy and have fewer visits, and, in any random cross section, actual completed family size (CHLDLIV) is likely to be highly correlated with desired family size, particularly when one has controlled for maternal age.

The estimated coefficients of MIC, CLINIC and OB/GYN in regressions (d) through (f) are consistent with regression (c) and the earlier regressions on INTERVAL and will not be discussed again at this point.

The coefficients of the various measures of the interval to the first visit (INTER1, INTER2, INTER3, INTER(350), INTER(350)2, INTERVAL) all demonstrate the very strong effect of the largely mother determined interval to the first visit on the actual number of visits. For example, the coefficients of the three dummy variables, INTER1, INTER2, INTER3 in regression (d) would indicate that other things equal women who begin care during their first trimester will have 9.1 visits, those who begin during their second trimester 7.9 visits, and those in the third 5.1 visits. The number of visits corresponding to each trimester is similar to the number of visits which would result from following the protocol for prenatal care as recommended by obstetricians. Thus, following this protocol, a mother would have approximately 12 visits if she began during the first trimester, 10 if she began in the second and 7 if she began in the third. Of course, these estimates would be associated with an uneventful pregnancy; pregnant mothers at high risk or suffering complications of pregnancy would be advised to have more visits. The persistent difference of between two and three visits between the recommended protocol and the estimated coefficients may have largely been the result of the restrictive upper bound of 11.04 placed on the dependent number of visits variable. This upper bound was necessitated by the manner in which the data was recorded.

Chapter VII - Estimation of Outcome Relationships

In the previous chapter, I investigated empirically the factors associated with the demand for and utilization of prenatal care. In this chapter, I investigate empirically the factors associated with different pregnancy outcomes, specifically birth weight, neonatal death, postneonatal death and infant death as a whole. In the analysis that follows, I shall particularly concentrate on attempting to determine the extent to which differences in the level of prenatal care utilized during a pregnancy contribute to differences in outcomes.

Birth weight

As discussed in Chapter IV, there are a number of good reasons for choosing birth weight as a measure of the success of a pregnancy. It is universally and fairly accurately measured and recorded. It is a good objective indicator of the condition of the child at birth, particularly as it is the best single predictor of subsequent infant mortality. Low birth weight has been found to be associated, among surviving infants, with mental and physical growth retardation during subsequent childhood years. It is a continuous variable which doesn't present special problems of estimation in a regression format. It would appear amenable to medical intervention, particularly through the control of the diet of pregnant women.

The results of regressing birth weight on a group of predetermined socio-economic and pregnancy experience variables as well as measures of prenatal care inputs are reported in Table 1.

In column (a) are reported the results of regressing birth weight

Table 1 (VII)

WGHT: BIRTH WEIGHT IN GRAMS

Coefficients of OLS Regression Estimates
Entire Sample - N=54,280
(t-statistic in parenthesis)

Independent Variables	(a)	(b)	(c)	(d)
AGEMOTH	26.369	21.544	14.638	7.253
	(7.80)**	(6.22)**	(4.55)**	(2.25)*
EDMOTH	-0.297	0.768	0.546	-0.190
	(0.33)	(0.86)	(0.65)	(0.23)
EDFATH	-0.616	0.732	0.863	-0.977
	(0.53)	(0.62)	(0.79)	(0.90)
LEGIT	99.285	102.593	82.145	56.826
	(14.54)**	(14.99)**	(12.9)**	(8.86)**
RACEN	116.110	-123.632	-91.234	-71.682
	(19.13)**	(20.21)**	(16.04)**	(12.39)**
FORMOTH	46.184	54.949	53.499	60.414
	(7.44)**	(8.79)**	(9.22)**	(10.42)**
PRMOTH	-31.148	-36.702	-28.397	-8.115
	(4.15)**	(4.89)**	(4.07)**	(1.16)
AGEMSQ	-0.371	-9.329	-0.204	-0.101
	(6.03)**	(5.26)**	(3.51)**	(1.74)
CHILDLIV		21.848	21.168	27.449
		(9.34)**	(9.75)**	(12.62)**
TOTLOSS		-36.672	-31.313	-32.912
		(9.15)**	(8.42)**	(8.90)**
FIRST		-8.526	-21.242	-23.762
		(1.31)	(3.52)**	(3.96)**
GESTN			-8.674	8.375
			(27.12)**	(86.96)**
INTER1				140.207
				(8.83)**
INTER2				-30.800
				(5.54)**

Table 1 (VII) Cont'd.

Independent Variables	(a)	(b)	(c)	(d)
INTER3				1.320 (0.19)
NUMVISIT				13.637 (14.67)**
$\bar{R}^2$	0.031	0.035	0.169	0.178
CONSTANT	2720.299	2771.136	462.781	470.766

*Statistically significant at 5% level

**Statistically significant at 1% level

on a limited number of non-pregnancy specific predetermined variables. As in previous regressions, the effect of mother's age is nonlinear - birth weight rises with maternal age up to age 35.5 and falls subsequently. This finding is consistent with the generally accepted notion that pregnancies are at greater risk as maternal age increases beyond 35 years. The coefficients on the group of characteristic dummy variables LEGIT, RACEN, FORMOTH and PRMOTH are all significant and consistent with the demand for care equations.* However, unlike the case of the previously reported demand for care equations these significant coefficients need not be regarded as indicators of our ignorance. Several of them, RACEN, PRMOTH and perhaps FORMOTH, may be regarded as indicating potential genetic or biological differences in physiological pregnancy production processes associated with racial or ethnic differences. Surprisingly, the coefficients on mother's and father's education are negative and insignificant. A priori, I would expect positive significant coefficients on these variables since (1) mother's education may be regarded as a measure of reproductive efficiency which should result in increased birth weights and (2) father's education has been regarded as a proxy for permanent income and has been shown to be related to the demand for and use of prenatal care and presumably other pregnancy inputs.

In regression (b), the experience variables CHILDLIV, TOTLOSS and FIRST are entered. Coefficients on these variables are significant or nearly so and the signs are as expected. The indication is that

*Note that regressions (a) through (c) may be interpreted alternatively as demand, production or outcome equations.

previous good experience as measured in CHILDLIV tends to repeat itself in the form of higher birth weights on subsequent pregnancies and that a history of pregnancy losses is associated with lower birth weights. Other things equal, therefore it would appear that families are somewhat justified in modifying their demand for care during a particular pregnancy based on their experience with prior pregnancies.

In regression (c), gestation age is entered into the equation. In a more complete empirical model of the pregnancy production process gestation age would be viewed as an intermediate outcome variable similar to birth weight; however, in the system we are estimating we are forced, by data limitations on our ability to identify specific relationships, to choose between estimating functions for either birth weight or gestation age and have choosen birth weight for reasons detailed more fully in Chapter 4. Here we note, based on regression (c), that there is a strong gestation period growth effect on birth weight - the fetus grows approximately 8.7 grams for each day of gestation.

In regression (d), the variables measuring prenatal care inputs are entered, INTER1, INTER2, INTER3, and NUMVISIT. We note that holding gestation age constant there is a 140 gram increase in birth weight associated with starting care during the first trimester. The gain associated with starting care during the second trimester falls to 110 grams and is also 110 grams for pregnancies where the initiation of care is delayed to the third trimester. Thus, the minimal increase in birth weight associated with any amount of prenatal care is 110 grams - a not inconsiderable amount when one considers that the traditional upper bound used to categorize infants as being of low weight is 2,500 grams.

The coefficient of NUMVISIT indicates that weight increases approximately 13.6 grams with each visit. Thus the total gain associated with a full complement of prenatal care comprising 12 visits and starting in the first trimester would be 303 grams as compared to a pregnancy without any care inputs. This increment in birth weight is equal to about 12% of the 2,500 low birth weight-high risk classification marker and about 10% of the mean birth weight for the entire sample.

It is worth comparing the value of the coefficients of the four maternal characteristic dummies (LEGIT, RACEN, FORMOTH, PRMOTH) in equation (a) with their value in regression (d) to determine whether the introduction of the experience, gestation age and prenatal care variables has an effect in reducing the otherwise apparently largely unexplained differences between these groups. The coefficient on LEGIT while still positive and significant is reduced by 44% primarily due to the importance of legitimacy status in determining the demand for care. The coefficient on RACEN is reduced by about 40% with half of that decrease attributable to the fact that blacks apparently have shorter gestation periods and the other half to the fact that they use fewer prenatal care inputs. Foreign born mothers apparently bear even heavier children than expected, particularly when account is taken of their lower utilization of prenatal care inputs. The most substantial effect is on the coefficient for Puerto Rican mothers. It drops from a highly significant -31 grams to an insignificant -8 grams. This would indicate that practically all of the birth weight differential that can be attributed to this ethnicity characteristic can be explained by differences in the utilization of prenatal care. Overall it would appear that a significant proportion of birth weight differentials that have been traditionally attributed to maternal characteristics such as race or to legitimacy status can be explained by variations in the demand for care among these different groups. That there are significant differentials remaining even after taking account of care inputs and the level of individual reproductive efficiency argues that there are still some important variations to be explained and that biological differences should be seriously considered as a possible explanation for certain of these differences.

Infant Deaths

In this section, I present the results of regressing a dichotomous dependent variable (1 = death) for various classifications of infant deaths on a group of predetermined socio-economic and pregnancy experience variables, measures of the level of prenatal care inputs, and intermediate outcome variables, such as birth weight, intended to account for the condition of the infant at birth. Regressions for three classes of outcomes are presented. In Table 2, the dichotomous dependent variable is one in the case of an infant death defined as the death of an infant born alive which occurs at up to one year of age. In Table 3, the dichotomous dependent variable is one if the death was a neonatal death, defined as the death of a live born infant within the first 28 days of life. In Table 4, the dichotomous dependent variable is one if the death occured during the post-neonatal period, i.e., between 28 days and one year of age. The sample on which regressions in Table 4 were run was restricted to only those infants who survived the first 28 days of life and hence were at risk during the post-neonatal period.

Table 2 (VII)

INFANT DEATH: Dichotomous Dependent Variable (Dead Infant = 1)
Entire Sample, N = 54,280
Coefficients of OLS regression Estimates
(t-statistic in parenthesis)

Independent Variables	(a)	(b)	(c)
MOTHAGE	-.0017 (2.04)*	-.0012 (1.51)	.0001 (.12)
MOTHED	-.0004 (2.00)*	-.0003 (1.63)	-.0002 (1.16)
FATHED	-.0002 (.68)	-.0002 (.58)	.00001 (.04)
RACEN	.0079 (5.45)	.0064 (4.19)*	-.0005 (.38)
FORMOTH	-.0024 (1.64)	-.0020 (1.35)	.0005 (.34)
PRMOTH	-.0015 (.87)	-.0031 (1.66)	-.0012 (.73)
CHILDLIV	.0003 (.51)	-.0004 (.63)	-.00004 (.08)
TOTLOSS	.0046 (4.86)**	.0047 (4.99)**	-.0007 (.84)
FIRST	-.0033 (2.15)	-.0026 (1.70)	-.0025 (1.85)
AGEMSQ	.00002 (1.62)	.00002 (1.27)	-.00001 (.54)
MIC		.0021 (1.63)	.0001 (.06)
INTER(350)		.0001 (9.60)**	-.00003 (2.54)*
NUMVISIT		-.0035 (15.15)**	.0004 (1.90)
NOVISIT		.0452 (11.51)**	.0094 (2.67)**

Table 2 (VII) Cont'd.

Independent Variables	(a)	(b)	(c)
ICU		.0007 (.59)	.0003 (.31)
WGHT			-.0005 (99.40)**
WGHTSQ			.0743 (91.26)**
GESTN			-.0009 (11.76)**
LEGIT	-.0073 (4.54)**	-.0044 (2.66)**	.0010 (.68)
CONSTANT	.055	.045	.916
$\bar{R}^2$	.003	.012	.215

*Statistically significant at the 5% level.

**Statistically significant at the 1% level.

As discussed in Chapter 4, neonatal deaths account for over 75% of all infant deaths, occur largely during the first several days of life in a hospital environment and appear to be associated with stresses on the infant associated with the prenatal environment and the birth process. On the other hand, post-neonatal deaths are largely attributable to different causes and appear to be significantly influenced by the home environment into which the infant is introduced after leaving the hospital. In the discussion that follows, I shall concentrate primarily on comparing and contrasting results for the two classes of death and not on infant death itself. This is partially because I am interested in testing the hypothesis advanced earlier (Chapter 4) about being able to evaluate the value of prenatal care per se by comparing its contribution to decreasing neonatal deaths with its effect on post-neonatal deaths and partially because, to the extent that different variables are alternatively responsible solely for either neonatal or post-neonatal deaths, their estimated coefficients in a regression for infant death will be biased toward zero. Therefore, the infant death regressions in Table 2 are presented as additional information for the interested reader but will not be discussed below.

Concentrating on the neonatal death regressions presented in Table 3, we note that, in the absence of information about care inputs and intermediate outcome measures (regression(a)), four predetermined variables appear to have a statistically significant effect on the probability of an infant's dying during the neonatal period and that the signs of these significant coefficients are all as might be expected, either because of predictions of the theoretical model or the results

Table 3 (VII)

NEONATAL DEATHS: Dichotomous Dependent Variable (1=Dead at less than 28 days old).

Coefficients of OLS Regression Estimates
(t-statistic in parenthesis)

Independent Variables	(a)	(b)	(c)
AGEMOTH	-.0004 (.57)	-.00005 (.06)	.0012 (1.96)*
EDMOTH	-.0004 (2.22)*	-.0003 (1.81)	-.0002 (1.31)
EDFATH	.0002 (.73)	.0002 (.79)	.0003 (1.65)
LEGIT	-.0065 (4.62)**	-.0040 (2.78)**	.0012 (.93)
RACEN	.0053 (4.19)**	.0045 (3.31)	-.0022 (1.88)
FORMOTH	-.0009 (.73)	-.0004 (.34)	.0019 (1.72)
PRMOTH	-.0004 (.28)	-.0014 (.88)	.0005 (.35)
AGEMSQ	.00001 (.40)	0	-.00002 (2.12)*
CHILDLIV	-.0004 (.73)	-.0009 (1.87)	-.0007 (1.54)
TOTLOSS	.0039 (4.77)**	.0041 (4.91)**	-.0012 (1.73)
FIRST	-.0021 (1.56)	-.0014 (1.08)	-.0014 (1.18)
MIC		.0009 (.81)	-.0010 (1.04)
INTER(350)		.0001 (10.32)**	-.00003 (3.09)**
NUMVISIT		-.0032 (15.74)**	.0006 (3.17)**

Table 3 (VII) Cont'd.

Independent Variables	(a)	(b)	(c)
NOVISIT		.0447 (12.99)**	.0099 (3.32)**
ICU		.0004 (.43)	.0001 (.12)
WGHT			-.0005 (115.87)**
WGHTSQ			.0737 (106.73)**
GESTN			-.0003 (12.74)**
CONSTANT	.0272	.0179	.873
$\bar{R}^2$	.002	.011	.266

*Statistically significant at the 5% level

**Statistically significant at the 1% level of the estimated demand and birth weight equations. Thus, mother's education has a small negative effect on the probability of death while a history of previous pregnancy losses is associated with an increased probability of death. In fact, the probability of death appears to increase 29% relative to the mean neonatal mortality rate for each previous pregnancy loss. Both these effects are consistent with viewing these variables as measures of reproductive efficiency. Illegitimacy and being black are associated with a higher probability of death during this period although as they are also associated with a much lower level of medical care inputs and low birth weight, it remains to be seen whether they have any independent effect.

In regression (b), five measures of medical inputs are included in the regression, MIC, INTER(350), NUMVISIT, NOVISIT, and ICU. Of these variables only the coefficients on INTER(350), NUMVISIT and NOVISIT are significant. The signs of the coefficients of number of visits (NUMVISIT) and the dummy variable standing for no prenatal care visits (NOVISIT) are consistent with regarding prenatal care visits as productive inputs into the pregnancy process. Thus each prenatal visit appears to reduce the probability of death by .003 or 22% of the mean probability while having no visits increases this probability by .045 or over 300 per cent.

The sign of the coefficient of INTER(350), a transformation of the interval to the first visit, appears at first reading to be contrary to expectations.* It would appear that this apparently perverse sign may

---

*Recall, INTER(350) equals 350 day minus the interval in days between the date of the mother's last menstrual period and her first prenatal care visit if she had at least one visit and zero otherwise.

result from the interaction of INTER(350) and the birth weight and gestation age variables which are not included in this regression. As we have seen in the results reported in Table 1, birth weight, which is the best predictor of infant survival and condition at birth, depends critically on the length of the gestation period. It is during this period that, other things equal, the fetus grows in size and develops facilities to cope with the external environment. Since INTER(350) equals zero if there have been no prenatal visits, and the NOVISIT coefficient accounts for the effect of having no visits, its coefficient can be viewed as a measure of the effect of lengthening the period to the first visit among those women who had visits. Consider the interpretation of INTER(350) among those women who delay the first visit until late in their pregnancy. Some of them will deliver before their first visit and INTER(350) for them will be zero - the effect of the outcome is measured in NOVISIT. On the other hand, those who start care late will have experienced a long gestation period by the time they start. If beginning care early in a pregnancy does not significantly increase gestation age, then INTER(350) may be regarded partially as a surrogate measure of perhaps otherwise high risk mothers who deliver at higher gestation ages when risk of adverse outcomes are reduced. Hence, long first prenatal visit intervals might in an inadequately specified model appear to be associated with reduced risk of neonatal mortality.*

---

*In the sample restricted only to those mothers who had at least one prenatal visit, the simple correlation between INTER(350) and GESTN is .12 - not large, but statistically significant at the 5% level.

The introduction of these care measurement variables also reduces the size of the coefficients associated with legitimacy and race. This is in agreement with the hypothesis advanced above that part of the substantial effect attributable to these variables is the result of differences in the utilization of prenatal care associated with these characteristics.

In regression (c), we introduce birth weight (WGHT and WGHTSQ) and gestation age (GEST). Not surprisingly, they are highly significant predictors of neonatal death. The probability of death declines with an increase in gestation age due to the continued time available for the fetus to mature before birth. The effect of birth weight is nonlinear. The probability of a neonatal death falls as weight rises and is a minimum at 3,400 grams. It begins to increase at weights beyond this point perhaps because of the increased risk of complications at delivery of extremely large infants.

With birth weight and gestation age in the regression, the coefficient of TOTLOSS looses significance as do both the coefficients of LEGIT and RACEN. The almost significant negative coefficient on RACEN is consistent with some observations that after standardizing for birth weight black infants have lower mortality rates.

Importantly, from the prospect of testing our hypothesis about the importance of prenatal care, all three prenatal care variables are significant. The signs of the coefficients of INTER(350) and NOVISIT are as expected, indicating that early care is better than late care and any care is better than none. The positive sign on NUMVISIT is somewhat unexpected, however, but may indicate that holding date of first visit, gestation age and birth weight constant, an increase in prenatal visits is associated with otherwise unaccounted for complications of pregnancy which result in marginally higher neonatal mortality rates.

Interestingly, the coefficient on ICU (baby born in a hospital with a neonatal intensive care unit) is insignificant despite reports of substantial success with these units (IOM, 1973). There are several reasons for believing that this finding is not at odds with other reports of substantial value added in ICU's. One is that a referral network has developed to shunt high risk mothers into hospitals with this special facility - hence, the high level of inputs in these units is balanced by the high degree of medical care needed by their patients. Second, a well developed postpartum transfer network exists within the City to move high risk infants immediately after birth to designated hospitals which have these facilities (the ICU variable refers to hospital of birth). Lastly, over 43% of all births in the City whether at high risk or not occur in hospitals with these ICU facilities as they tend to be located in hospitals with large obstetrical services.

Turning our attention to the set of POSTNEONATAL death regressions reported in Table 4, we note as expected, that different variables are associated with postneonatal deaths than neonatal deaths.

In particular, we note that postneonatal death is consistently affected by a group of predetermined variables included in regression (a) whose significance is not materially affected by the inclusion of prenatal care measures in regression (b) or the condition of the infant at birth as included in (c). Moreover, most of these effects are difficult to explain within the context of our model. Blacks have higher than

Table 4 (VII)

POSTNEONATAL DEATHS: Dichotomous Dependent Variable (1=Infant dead between 28 days and 1 year of age)

Coefficients of OLS Regression Estimates

Sample Restricted to Births Surviving at least 28 days--N=53,592

Independent Variables	(a)	(b)	(c)	Mean Standard Deviation
AGEMOTH	-.0013	-.0011	-.0011	25.17
	(3.19)**	(2.88)**	(2.75)**	5.5.
EDMOTH	-.00001	-.00001	-.00001	10.95
	(.14)	(.10)	(.10)	3.67
EDFATH	-.0004	-.0003	-.0003	11.90
	(2.69)**	(2.51)*	(2.50)	2.60
LEGIT	-.0008	-.0003	-.0001	.79
	(1.03)	(.37)	(.15)	.41
RACEN	.0026	.0019	.0016	.29
	(3.68)**	(2.54)*	(2.19)*	.45
FORMOTH	-.0015	-.0017	-.0014	.22
	(2.08)*	(2.28)*	(1.96)*	.42
PRMOTH	-.0011	-.0018	-.0017	.16
	(1.30)	(1.98)	(1.88)	.38
AGEMSQ	.00002	.00002	.00002	663.68
	(2.61)**	(2.37)*	(2.22)*	300.12
CHILDLIV	.0006	.0005	.0006	1.11
	(2.36)*	(1.95)*	(2.14)*	1.48
TOTLOSS	.0007	.0007	.0005	.19
	(1.44)	(1.45)	(1.07)	.63
FIRST	-.0012	-.0012	-.0012	.41
	(1.62)	(1.57)	(1.60)	.49
MIC		.0012	.0011	.37
		(1.85)	(1.67)	.48
INTER(350)		.00004	-.00001	234.35
		(1.05)	(.18)	71.71

Table 4 (VII) Cont'd.

Independent Variables	(a)	(b)	(c)	Mean Standard Deviation
NUMVISIT		-.0003 (2.60)**	0.0002 (1.49)	8.02 3.28
NOVISIT		-.0033 (.93)	.0003 (.08)	.03 .17
ICU		.0002 (.39)	.0002 (.38)	.43 .50
WGHT			-.00002 (5.47)**	3195.64 533.02
WGHTSQ			.0027 (4.75)**	10.50 3.34
GESTN			-.00003 (2.22)*	278.91 25.77
CONSTANT	.0285	.0279	.0673	
$\bar{R}^2$	.002	.002	.003	

*Statistically significant at the 5% level

**Statistically significant at the 1% level expected postneonatal death rates but children of foreign born mothers lower than expected postneonatal death rates. The probability of death during this period declines with mother's age reaching a low point at 27.5 years and rising thereafter. The reason for this particular pattern is unclear.

Two findings that appear consistent with the model advanced earlier are the finding that father's education, a proxy for permanent income, is negatively correlated with deaths at this age. This is consistent with viewing children as a normal good, the demand for which increases with income. Moreover, the finding that the number of live children in the home tends to increase mortality is consistent with either the notion that increasing family size strains the resources available per child or that people who plan on large families consciously plan on spending less per child - alternatively one could argue that in large families additional children are "unwanted".

Of particular importance from the view point of testing our hypothesis about the actual value of prenatal care inputs in producing successful pregnancies is the finding that the coefficients of most care variables which were significant in the regressions for neonatal death are not significant in the estimated postneonatal death regressions. In particular, in regression (c), we note that although birth weight and gestation age continue to importantly influence survival during the post-neonatal period (as well they might as indices of the quality of the infant at birth), INTER(350), NOVISIT, and NUMVISIT are not statistically significant.

Chapter VIII - Full Information Maximum Likelihood and Classical Least Squares Estimations Compared

One of the most important issues surrounding the finding in the Institute of Medicine (1973) and other studies of a positive effect on infant survival of "adequate" prenatal care is whether this finding is due to a "third" variable. That is, the amount of prenatal care demanded during an individual pregnancy while not productive in itself may serve as a proxy for other variables which are important and statistically may appear to be a good predictor of infant survival. Candidates for such third variables include the degree of wantedness of the child, other productive inputs such as good nutrition which cannot be directly measured in vital statistics data but should correlate highly with prenatal care or even the ability of the parents to effectively use the medical care system when necessary.

In Chapter IV, I argue that if prenatal care is serving only as a proxy for some third variable such as "wantedness" or nutrition, it should probably serve equally well as a predictor of infant survival during the neonatal and postneonatal period. However, since the value of medical care received before and during birth should be most significant during the neonatal period, a finding that prenatal care significantly improved chances of survival during the neonatal period, particularly holding birth weight and gestation age constant and that prenatal care was not a significant factor during the postneonatal period, would serve as strong evidence that prenatal care per se was indeed an input into infant health.

At the conclusion of the previous chapter (Chapter VII), I note that, based on the results of OLS regression estimates of functions with either neonatal or postneonatal death as dependent variables, it would appear that prenatal care is an important input into the production of healthy infants as measured by their ability to survive the neonatal period. This conclusion is tempered somewhat by the fact that it is based on OLS estimating techniques in a situation where the dependent variable is dichotomous. As discussed in Chapter V, many econometricians have expressed objections to using OLS estimation for these relationships. Accordingly, estimates of the relationships for neonatal death and postneonatal death as well as the dichotomous visit/no visit demand equation were attempted using a full information maximum likelihood logit estimation procedure.

Unfortunately, such FIML estimators can only be calculated using an iterative procedure whose cost increases with the number of observations, the number of independent variables and the number of iterations required to reach convergence of the likelihood function. It was physically impossible, as well as potentially enormously expensive, to perform these estimations on the entire sample of over 54,000 births for the 15 to 20 independent variables under investigation. Accordingly, several randomly drawn subsamples of manageable size were constructed for these estimates. In all instances, these subsamples consisted of a random sample of 50% of the population with the attribute (NOVISIT, NEONATAL DEATH, POSTNEONATAL DEATH) and a 4% sample of the remaining population. Individual observations were weighted inversely proportionately to their sampling fraction so that subsample means and functional estimates would reflect the actual population from which the sample was drawn.

FIML logit estimates were run for each dependent variable on the appropriate subsample and compared with similarly weighted OLS regressions on each subsample and unweighted OLS regression estimates on the entire population. The results of these additional estimates and a comparison of estimated relationships using different forms of estimation are presented in Table 1 for neonatal death, in Table 2 for postneonatal death and in Table 3 for NOVISIT.

Taken as a whole, the comparisons are somewhat disquieting. There is little agreement between the FIML estimates and the OLS estimates for the entire sample. Considering neonatal death first, we note that in regression (c) of Table 1, only birth weight, as measured by WGHT and WGHTSQ, appears to be a significant predictor of neonatal death. Even the coefficient on gestation age is not significant. In the OLS regression on the entire sample, race, mother's age and the measures of prenatal care as well as gestation age all have significant coefficients whose signs are as predicted by theoretical and other considerations. Interestingly, in regression (a) (a FIML estimate) of Table 1 no variable appears with a significant coefficient although significant relationships have been reported between race or legitimacy and neonatal death based on contingency tables. In regression (b) the prenatal care variables are significant with signs consistent with those reported in the OLS regression in regression (b) of Table 2, Chapter VII. However, these effects appear to wash out when birth weight is included in regression (c).

Comparing the OLS regression on the sample regression (d), with the OLS regression on the entire population and the FIML estimate on the sample, we note that for the three variables WGHT, WGHTSQ and GESTN,

Table 1 (VIII)

NEONATAL DEATH: Dichotomous Dependent Variable: Full Information on Maximum Likelihood Logit Estimates and OLS Estimates Compared

(∂neonatal death $/\partial X$ in [ ] brackets)

(t-values in parenthesis)

Independent Variables	FIML Logit (Subsample)[a] (a)	(b)	(c)	OLS Subsample[a] (d)	OLS Population (e)
MOTHAGE	.0771 (.32) [.001]	.1030 (.42) [.001]	.3574 (1.20) [.005]	.0043 (1.57)	.0015 (2.40)*
MOTHED	-.0417 (-.76) [-.0006]	-.0338 (-.60) [-.0005]	-.0110 (-.15) [-.0001]	.0001 (.14)	-.002 (1.55)
FATHED	-.0235 (-.30) [-.0003]	-.0237 (-.30) [-.0003]	-.0473 (-.44) [-.0006]	-.0002 (.24)	.0003 (1.63)
RACEN	.3012 (.81) [.004]	.2856 (.75) [.004]	-.3317 (-.64) [-.004]	-.0059 (1.29)	-.0026 -.0026 (2.46)*
LEGIT	-.2756 (-.64) [-.004]	-.1291 (-.29) [-.002]	.1687 (.28) [.002]	-.0014 (.93)	.0015 (1.25)
CHILDLIV	-.0567 (-.43) [-.0008]	-.1217 (-.88) [-.002]	-.1855 (-1.02) [-.002]	-.0014 (.91)	-.0005 (1.43)
TLOSS	.3168 (1.46) [.004]	.3197 (1.40) [.004]	-.1855 (-.47) [-.002]	-.0040 (1.10)	-.0011 (1.60)
AGEMSQ	-.0014 (-.32) [-.0002]	-.0016 (-.35) [-.0002]	-.0055 (-.99) [-.00007]	-.00007 (1.45)	-.00003 (2.48)*
ICU		.0882 (.26) [.001]	.0899 (.20) [.001]	.0021 (.53)	.0001 (.08)

NOENATAL DEATH: Dichotomous Dependent Variable - Cont'd.

Independent Variable	FIML Logit (Subsample)[a] (a)	(b)	(c)	OLS Subsample[a] (d)	OLS Population (e)
INTER(360)		-.0093	-.0005	-.0001	-.00003
		(2.44)* [.001]	(-.10) [-.00001]	(1.27)	(3.19)**
NUMVISIT		-.2693	-.0132	.0006	.0006
		(-3.73)** [-.004]	(-.14) [-.0002]	(.70)	(3.18)**
NOVISIT		1.8590	1.0963	.0070	.0098
		(2.12)* [.024]	(.88) [.015]	(.49)	(3.28)**
WGHTGMS			-.0062	-.0005	-.0005
			(-4.83)** [-.00008]	(26.43)**	(115.91)**
WGHTSQ			.8664	.0798	.0737
			(3.54)** [.012]	(24.60)**	(106.79)**
GESTN			-.1280	-.0003	-.0003
			(-1.56) [-.0002]	(3.35)**	(12.80)**
Constant	-4.49	-4.96	4.42		
$\bar{R}^2$				.270	.266
χ^2	4.78	26.37**	165.71**		
N	2656	2656	2656	2656	54,280

*Significant at the 5% level.

**Significant at the 1% level.

[a]Weighted regression based on a random sample of 50% of all neonatal deaths and 4% of all other live births with weights inversely proportional to the sampling fraction.

the coefficients in the OLS regression on the sample are identical with the OLS estimates on the entire population. However, the lack of statistically significant coefficients in regression (d) on the maternal age, race, and care variables is more in keeping with the results of the FIML estimates in (c) and raises the question as to whether the estimates based on the sample are biased by factors peculiar to the sample chosen.

The result for POSTNEONATAL DEATH (Table 2) are also inconsistent. In none of the FIML estimates regressions, (a) through (c) is any coefficient statistically significant; none of the χ^2 statistics associated with the entire estimated equation are significant and in the case of equation (c) the likelihood function did not converge. The lack of statistical significance is also found in the weighted OLS regression, (d), run on this same sample for which the overall F statistic of .50 is not statistically significant at the 5% level. The OLS estimate based on the entire sample again contains many statistically significant coefficients whose signs are consistent with a priori predictions and contingency table analyses by others (Shapiro, Schlesinger, and Nesbitt, 1968).

This pattern of differences in between estimated relationships of the same dependent variable obtained by different estimating procedures is repeated in the case of NOVISIT (Table 3). Once again, OLS subsample estimates (regression (d)) bear more resemblance to FIML estimates (regression(c)) than to OLS estimates on the entire population, and very few variables have significant or consistent estimated coefficients in the estimates based on the subsample.

Table 2 (VIII)

POSTNEONATAL DEATH: Dichotomous Dependent Variable: Full Information Maximum Likelihood Logit Estimates and OLS Estimates Compared

(∂neonatal death/∂X in [ ] brackets)

(t-values in parenthesis)

Independent Variables	FIML Logit (Subsample)[a] (a)	(b)	(c)	OLS Subsample[a] (d)	OLS Population (e)
AGEMOTH	-.1912 (-.47) [-.0008]	-.1854 (-.45) [-.0008]	-.1194 (-.24) [-.0005]	-.0009 (.45)	-.0012 (2.97)**
EDMOTH	-.0275 (-.29) [-.0001]	-.0308 (-.27) [-.0001]	-.0274 (-.24) [-.0001]	-.0001 (.13)	.00003 (.24)
EDFATH	-.0802 (-.79) [-.0004]	-.1058 (-.70) [-.0004]	-.1101 (-.72) [-.0005]	-.0003 (.48)	-.0003 (2.21)*
LEGIT	-.0779 (-.11) [-.0004]	.0490 (-.06) [-.0002]	.0027 (.003) [.00001]	.0002 (.06)	-.0002 (.27)
RACEN	.8284 (1.10) [.003]	.7598 (1.07) [.003]	.6952 (.97) [.003]	.0029 (.90)	.0024 (3.55)**
CHILDLIV	.1484 (.55) [.0005]	.1029 (.44) [.0004]	.1029 (.44) [.004]	.0001 (.08)	.0006 (2.35)*
TOTLOSS	.3485 (.55) [.001]	.2434 (.53) [.001]	.1823 (.40) [.0007]	.0008 (.31)	.0005 (1.03)

Table 2 (VIII)

POSTNEONATAL DEATH: Dichotomous Dependent Variable - Cont'd.

Indepdendent Variable	FIML Logit (Subsample)[a] (a)	(b)	(c)	OLS Subsample[a] (d)	OLS Population (e)
AGEMSQ	.0030 (.36)	.00275 (.36)	.0016 (.22)	.00001 (.38)	.00002 (2.38)*
INTER(350)		.0021 (.29) [.00001]	.0003 (.04) [.00001]	.0000 (.03)	-.00002 (.20)
NUMVISIT		-.0743 (-.55) [-.0003]	-.0351 (-.24) [-.0001]	-.0001 (.21)	-.0001 (1.40)
NOVISIT		.6747 (.38) [.003]	.5055 (.28) [.002]	.0046 (.45)	.0004 (.10)
ICU		.0407 (.06) [.002]	.0763 (.12) [.0003]	.0003 (.11)	.0002 (.37)
WGHT			.0014 (-.51) [-.00001]	-.00002 (1.15)	-.00002 (5.53)**
WGHTSQ			.1886 (.40) [.0008]	.0031 (1.06)	.0027 (4.80)**

Table 2 (VIII)

POSTNEONATAL DEATH: Dichotomous Dependent Variable - Cont'd.

Independent Variable	FIML Logit (Subsample)[a] (a)	(b)	(c)	OLS Subsample[a] (d)	OLS Population (e)
GESTN			.0055 (-.27) [-.00002]	-.00003 (.40)	-.00003 (2.23)*
Constant	-1.4205 (-.27)	-1.5629 (-.29)	1.6467 (.98)	.0653	.0677
R^2	-	-	-	.0035	.0034
F	-	-	-	.50	11.45**
N	2288	2288	2288	2288	53592
X^2	4.39	4.98*	5.97*	-	-

Table 3 (VIII)

NOVISIT: Dichotomous Dependent Variable: Full Information Maximum Likelihood Logit Estimates and OLS Estimates Compared

(∂ NOVISIT /∂X in brackets [ ])

(t-statistic in parenthesis)

	FIML Logit (Subsample)[a]			OLS Subsample[a]	OLS Population
Independent Variable	(a)	(b)	(c)	(d)	(e)
AGEMOTH	-.0410	-.1895	-.1831	-.0073	-.0068
	(-1.96)	(-1.30)	(-1.24)	(1.55)	(6.28)**
	[-.001]	[-.006]	[-.006]		
EDMOTH	-.0202	-.0123	-.0104	-.00002	-.0006
	(-.77)	(-.31)	(-.26)	(.01)	(2.20)*
	[-.001]	[-.0004]	[-.0003]		
EDFATH	-.1147	-.0886	-.0927	-.0018	-.0015
	(-2.33)	(-1.76)	(-1.82)	(1.06)	(4.01)**
	[-.004]*	[-.003]	[-.003]		
RACEN	.2390	.0927	.0540	-.0031	.0052
	(.83)	(.31)	(.18)	(.34)	(2.57)*
	[-.008]	[-.003]	[.002]		
LEGIT	-1.1883	-1.089	-1.1114	-.0473	-.0399
	(-4.56)	(-4.05)	(-4.07)	(4.78)**	(18.69)**
	[.038]**	[-.035]**	[-.036]**		
PRMOTH	.6454	.5481	.5393	.0238	.0176
	(2.14)	(1.77)	(1.71)	(2.08)*	(7.29)
	[.021]*	[.018]	[.017]		
CHILDLIV		.0304	.0880	.0032	.0107
		(.09)	(.24)	(.25)	(4.71)**
		[.001]	[.003]		

NOVISIT: Dichotomous Dependent Variable - Cont'd.

Independent Variable	FIML Logit (Subsample)[a]			OLS Subsample[a]	OLS Population
	(a)	(b)	(c)	(d)	(e)
%LOSS		-.6922	-.0897	-.0026	.0078
		(-.64)	(-.07)	(.08)	(1.14)
		[-.022]	[-.003]		
AGEMSQ		.0018	.00161	.0001	.0001
		(.64)	(.58)	(.98)	(4.50)**
		[.0001]	[.0001]		
TBO		.1886	.1393	.0051	-.0028
		(.61)	(.43)	(.46)	(1.44)
		[.006]	[.004]		
FIRST		-.2979	-.2938	-.0068	-.0020
		(-.90)	(-.89)	(.67)	(.92)
		[.010]	[-.009]		
MIC			.1263	.0034	.0092
			(.48)	(.38)	(4.62)**
			[.004]		
LAST			-.7653	-.0170	-.0097
			(-1.01)	(.85)	(2.25)*
			[-.024]		
CLINIC			-.0002	-.00001	-.0001
			(-.84)	(.88)	(3.99)**
			[-.0001]		

NOVISIT: Dichotomous Dependent Variable - Cont'd.

Independent Variable	FIML Logit (Subsample)[a] (a)	(b)	(c)	OLS Subsample[a] (d)	OLS Population (e)
Constant	-.2709 (-.34)	1.6023 (.83)	1.6366 (.83)	.2109	.1872
R^2				.026	.028
N	2593	2593		2593	54.280
X^2	61.64**	72.96**	74.87**		

*Significant at the 5% level.

**Significant at the 1% level.

[a]Weighted regression based on a random sample of 25% of births with no prenatal visits and 4% of births with weights inversely proportional to the sampling fraction.

Overall, it is difficult to know how to interpret these findings. The FIML estimation program utilized in these experiments was developed to be used on much smaller bodies of data than were available in this study. Although advocates of this FIML logit approach have marshalled sophisticated mathematical arguments to support its use rather than classical OLS in the case of dichotomous dependent variables, I know of no study which either through calculation or Monte Carlo techniques has demonstrated the added value attributable to the FIML approach when data sets are very large. Moreover, it is not clear that it is desireable to substantially reduce the size of large data sets so that FIML iterative estimation procedures can be applied rather than use some form of OLS on the entire population. In fact, in discussing whether one should attempt to apply generalized least squares rather than OLS to at least adjust for heterscedasticity, Feldstein (1966) notes that the "gain in efficiency does not justify the additional calculations" (his sample is nearly 17,000) but gives no evidence for this statement.

Before closing this section on this completely uncertain note, I should like to address the question as to whether peculiarities in the subsample could have been responsible for these inconsistent results. There doesn't seem to be any reason a priori to assume that the sampling procedure was at fault. In fact, comparison of the actual coefficients in the OLS regressions on the subsamples and the entire population indicate frequent similarities between the size and sign of the estimated coefficients even when these coefficients are not statistically significant in the regressions on the smaller sample. Moreover, it should be noted that the technique of over sampling observations with the rare outcomes such as was done in constructing these samples is fairly common practice in infant mortality studies. In fact, Feldstein (1966) uses data from a British Perinatal Mortality Survey (Butler and Bonham, 1963) wherein data on live births was obtained for births occurring during a single week while perinatal death observations included deaths occurring during a three month period - the data was then weighted inversely proportionately to the sampling fractions. Similar procedures are followed in constructing national U.S. infant death surveys.

Chapter IX - Conclusion

We began this study by noting that infant mortality rates have traditionally been used as indices of the health status of designated population groups and that because the U.S. had in recent years lagged significantly behind other developed nations in the reduction of infant mortality, concern had been voiced about the U.S. health care system and policies recommended to increase resource outputs in an effort to reduce infant mortality in this country.

In Chapter 2, it is demonstrated that measured infant mortality rates depend not only on health status but also on the demand for children and the extent to which biologically ill equiped mothers attempt to replace unsuccessful pregnancies by repeated conceptions. In Chapter 3, we broaden the family's decision possibilities and consider not only the reasons why people might continue to have additional pregnancies in the face of a record of pregnancy losses but also how decisions on prenatal care and other pregnancy inputs might be determined rationally by family income/experience with previous pregnancies, completed family size and the degree of "wantedness" of the child. These factors are explored empirically in Chapters VI through VIII.

Beginning at the end let us consider the results of the experiment proposed in discussing the empirical formulation of the model that if prenatal care inputs favorably effect neonatal survival and not postneonatal survival then we shall consider this as evidence that care "matters". The evidence of the OLS estimates does bear this hypothesis out although the mechanism through which care works remains unclear. The estimated coefficients support the notion that some care and probably early care is better than no care but it is difficult to determine what level of care is optimal and how should it be delivered. The birth weight regressions suggest that care also significantly and substantially increases birth weight - although in this case the possibility of a proxy such as maternal nutrition should be kept in mind. Additional support for the value of care is found in the fact that inclusion of care variables in the regression equations substantially reduces differentials in infant death and birth weight attributable to racial or ethnic differences or legitimacy status.

The estimated demand equations to the extent to which they are consistent with the theoretical model suggest that decisions about seeking care are approached rationally by individuals. Families with a history of unsuccessful pregnancies seek more care than those who have experienced more success in previous pregnancies. It is difficult to distinguish the effect of learning from experience from a "wantedness" effect - although it would appear that advocates of hypothetical "wantedness" effects will have to be more careful about defining their terms before a meaningful test can be made of that hypothesis.

If one accepts the conclusion that prenatal care can be of substantial value in improving infant survival, it is distressing to note the large differentials in the demand for care observed among different ethnic and racial groups which cannot be explained away by other factors. Thus the finding that black and Puerto Rican mothers receive substantially less care, primarily because they start later, is of concern because they would probably be less likely to receive a full complement of care anyway based on their other characteristics. The same is particularly true of illegitimate mothers who as a group are least likely to receive any care and among those who receive care, receive the least care. The rising trend of illegitimacy in this country over the last two decades argues that this problem may become more serious in the future.

Lastly, like the negative utilization effect measured for blacks, Puerto Ricans, and illegitimate mothers, the negative significant signs on the Maternal and Infant Care Project variables give cause to reassess these programs. These programs have been established in areas where the characteristics of the population would have resulted in very low levels of care utilization anyway. One would have thought that at the least the MIC programs would have provided these high risk mothers with a level of care equivalent to that obtainable by similar women elsewhere in the city. Before one can fairly evaluate these programs, an attempt would have to be made to review the resources these programs have at their disposal and balance the cost of individual projects against the benefits that they may provide to their targeted populations.

Two closing caveats are in order. One is that the empirical work in this study is based on data from the period immediately before the legalization of abortion in New York City and subsequently nationally. It would appear that validation of many of the empirical results using post-abortion data would be in order. In fact, the impact of abortion on the demand for prenatal care and infant survival would be worthy of study in itself and might help shed some light on the debate concerning the desirability of legalized abortion. Whether or not the theoretical model needs to be rethought is unclear; however, I would expect that to the extent to which post legalization data reflected decisions about pregnancy and care that were more predictable (and hence contained less noise), a model based on rational utility maximization should be of more utility in explaining behavior in this often emotionally charged area.

The second caveat is concerned with the failure of the FIML logit estimates of the outcome equations to agree with the OLS estimates. Most of the preceding discussion is based on the tacit assumption that the experiment performed on the prenatal care variables as reflected in the OLS estimations is valid. If one examines the FIML estimations, one finds little support for the conclusion that prenatal care has value per se, although there is nothing in those results to contradict this result. Aside from drawing attention to the failure of these two estimation techniques to yield results consistent with one another, it is beyond the scope of this study to examine the reasons for this lack of agreement. Recently, econometricians (Nerlove and Press, 1973) have come out strongly for FIML logit estimators an it appears that economists are increasingly going to have to be concerned with dichotomous or even polytomous dependent variables (Quigley, 1976). At the same time, economists have increased access to and will probably make increased use of large microdata sets. The use of iterative FIML estimation techniques on these data sets is usually expensive if not prohibitive. Increasingly, econometricians who have demonstrated the theoretical superiority of FIML techniques are going to have to deal with the economic question as to the circumstances under which the improvement in the results justify the added cost of these techniques. Of course, this is but a variation of the question basic to economics as a discipline.

Appendix A: Relationship Between the Distribution of Reproducive Efficiency and Measured Infant Mortality Rates

In this appendix, I shall demonstrate by simple examples the mathematical difficulties encountered in attempting to predict the expected relationship between $E(p|c)$ and the moments of the distribution of p_i in a given population group. In particular, under several very limiting assumptions, it is possible to predict the direction of the relationship between $E(p_i)$ and $E(p|c)$ but the same does not hold for the relationship between the variance of p, σ_p^2 and $E(p|c)$.

<u>$E(p_i)$ and $E(p|c)$</u>

For certain "smooth" or "regular" movements in $E(p_i)$, it is possible to say something about the expected change in $E(p|c)$. Consider the following simple example:

Let us approximate $E(p_i)$ by

$$\mu_p \approx \sum p_i f_i^{*} \tag{1}$$

and $E(p|c)$ by

$$E(p|c) \approx \frac{1}{\sum \frac{f_i}{p_i}} \tag{2}$$

We do not lose any generality by assuming all p_i's $(0 \leq p_i \leq 1)$ fixed and that μ_p changes by varying the f_i's. We can increase μ_p as follows:

increase f_k by Δ so that $f_k = f_k + \Delta$

now since $\sum f_i = 1$ there must be at least one $f_j = f_j - \Delta$

when $p_j < p_k$, μ_p will increase by $\Delta(p_k - p_j)$.

Now, in the case of $E(p|c)$, $\sum \frac{f_i}{p_i}$ will change by $\frac{-\Delta}{p_j} + \frac{\Delta}{p_k}$

but $p_j < p_k$, therefore

$$\Delta\left(\frac{1}{p_j} - \frac{1}{p_j}\right) < 0 \tag{3}$$

---

*Throughout this appendix all sums are over all i such that $(0 \leq p_i \leq 1)$ and will be denoted simply by $\sum$.

and $\sum \frac{f_i}{p_i}$ will decline but $E(p|c) = \frac{1}{\sum \frac{f_i}{p_i}}$ will increase. If all changes in $E(p_i)$ are of this nature then $E(p_i)$ and $E(p|c)$ will be positively correlated.

σ_p^2 and E(p c)

Unfortunately, even the simplest change in the variance of p will not yield predictable changes in $E(p|c)$. Consider the following example of a change in σ_p^2 when $E(p_i) = \mu_p$ is held constant. By definition,

$$\sigma_p^2 = \left(\sum p_i^2 f_i\right) - \mu_p^2 \tag{1}$$

but μ_p is held constant and the p_i's are fixed $(0 \leq p_i \leq 1)$ so σ_p^2 can only change by varying the f_i's as above such that

$$\Delta(p_k^2) - \Delta(p_j^2) > 0 \tag{2}$$

where $p_k > p_j$, however, the situation is not as straightforward as previously since $\sum f_i = 1$ and μ_p is being held constant.

The simplest way to satisfy these two conditions is to partition the change in the f_i's, Δ , as follows

$$\Delta = a\Delta + (1-a)\Delta \tag{3}$$

(so that $\sum f_i = 1$ is maintained) and

$$\Delta p_y = a\Delta p_x + (1-a)\Delta p_z \tag{4}$$

where $p_x < p_y < p_z$ (so that μ_p is maintained constant) therefore if

σ_p^2 is to increase when $f_y = \tilde{f}_y + \Delta$ then

$$p_y^2 < ap_x^2 + (1-a)p_z^2 \; . \tag{5}$$

What can we say, in this case about $\sum \frac{f_i}{p_i}$ the denominator of $E(p|c)$? We note first that the condition that μ_p is constant, (4), translates into a change in $\sum \frac{f_i}{p_i}$ consisting of $+ \frac{\Delta}{p_y}$ and $-\left(\frac{a\Delta}{p_x} + \frac{(a-1)\Delta}{p_z}\right)$ dividing through by Δ and gathering terms only yields

$$\text{Change in } \left(\sum \frac{f_i}{p_i}\right) = \frac{1}{p_y} - \frac{ap_z + (1-a)p_z}{p_z p_x} \tag{6}$$

which could result in $\sum \frac{f_i}{p_i}$ either increasing, decreasing or remaining unchanged. The result is ambiguous because

(a) $p_y > p_z p_x$ (note $p_y > p_x$ and $1 \geq p_z$, so $p_y \cdot 1 = p_y > p_x p_z$) and

(b) $1 > ap_z + (1-a)p_x$ (note that even if we allow $p_z = 1$, its largest possible value this inequality would still hold for if $ap_x + (1-a)p_x = 1$ and $p_z = 1$, this implies $a + p_x(1-a) = 1$ or $1 - p_x = a(1-p_x)$ which is true only if $a = 1$ but $a < 1$).

Nor does (5) impose any constraint on the sign of (6). Note that (4) and (5) together require only that

$$p_x^2 + p_z^2 > p_x p_z \; . \tag{7}$$

Now if we let (6) equal zero (no change in $E(p|c)$) then

$$\frac{1}{p_y} = \frac{ap_z + (1-a)p_x}{p_z p_x} \quad \text{or} \tag{8}$$

$$p_z p_x = p_y(ap_z + (1-a)p_x) \tag{9}$$

substituting from (4) for p_y yields

$$p_x p_x = (ap_x + (1-a)p_z)(ap_z + (1-a)p_x) \tag{10}$$

multiplying out both sides of (10) and gathering like terms yields

$$2p_z p_x = p_x{}^2 + p_z{}^2 \tag{11}$$

which is perfectly compatible with (6). Moreover, if we changed the equal sign in (8) to $>$ or $<$ a similar change would occur in (11). In either case this would be compatible with (6). So that it would appear that the relationship between $E(p|c)$ and $\sigma_p{}^2$ would be largely unpredictable except for very specifically defined changes in the underlying distribution of p.

Empirically, this lack of predictability could easily result in a failure to uncover a statistically acceptable relationship between measured mortality rates and the moments of $f(p_i)$. Consider the above example regarding $\sigma_p{}^2$ and $E(p|c)$. If we examine the distributions of p for two population groups, say F_1 and F_2 where $\mu_1 = \mu_2$ and $\sigma^2_{p1} \neq \sigma^2_{p_2}$ then even if we could decompose all differences between F_1 and F_2 as above in (5), it is conceivable that some of the differences could reduce $E(p|c)$ and some increase $E(p|c)$ for distribution F_2. Therefore the net measureable effect on $E(p|c)$ of moving from distribution F_1 to distribution F_2 could well be zero, although individuals may be behaving in accordance with the implications of hypothesized replacement behavior.

Appendix B: Sources of Data

Primary data for the regression estimates of the demand for care and pregnancy outcome were contained on unedited data tape listings of birth and linked infant death-birth certificates supplied to me by the New York City Department of Health. The data on the tapes represented all births re gistered in New York City in 1970 on a single tape and on two additional tapes, linked birth and death records for deaths occurring in either 1970 or 1971. Deaths occuring in 1971 were required because of the possibility that infants born in 1970 could have died in 1971. These data sets were merged to create a single birth-death tape representing all births to New York City residents during January to June 1970 and all infant deaths corresponding to these births occurring within a year of birth (through June 1971).

Data on hospitals with neonatal intensive care units was obtained by telephone inquiry of individual hospitals who had births in 1970 recorded on the tape. Information on clinic hours and location as well as about MIC project facilities was obtained from the New York City Bureau of Maternity Services and Family Planning. Information about the number of obstetrician/gynecologists practicing in a given Health District was obtained from a special unpublished survey provided by the Health and Hospital Planning Council of Southern New York, Inc.

Bibliography

Becker, G.S. 1960. An Economic Analysis of Fertility. In Demographic and Economic Change in Developed Countries. Princeton: Princeton University Press.

____________ 1965. A Theory of the allocation of time. Econ. J. 75: 493-517.

____________ 1971. Economic Theory. New York: Knopf.

Becker, G.S. and Lewis, H.F. 1972. Interaction Between Quantity and Quality of Children. In Economics of the Family, ed. T.W. Schultz. Chicago: University of Chicago Press.

Ben-Porath, Y. 1973. On Child Traits and the Choice of Family Size. M. Falls Institute, Discussion Paper No. 731.

Ben-Porath, Y. and Welch, F. 1973. Chance, Child Traits and the Choice of Family Size. R-1117-NIH/FR. Santa Monica: Rand Corporation.

Ben-Porath, Y. and Welch, F. 1976. Do sex preferences really matter. Quar. J. Econ. XC:285-307.

Billewicz, W.Z. 1973. Some implications of self-selection for pregnancy. Brit. J. Prev. & Soc. Med. 27:49-52.

Butler, N.R. and Bonham, D.S. 1963. Perinatal Mortality. Edinburgh and London: E. and S. Livingstone Ltd.

Chard, T. 1974. The fetus at risk. The Lancet October 12, 1974:880-882.

Chase, H.C. 1967. Registration completeness and international comparisons of infant mortality. Demography 6:425-433.

____________ 1972. The position of the United States in international comparison of health status. Amer. J. Public Health 62:581-589.

DeTray, D. 1972. Child Quality and the Demand for Children. In Economics of the Family, ed. T.W. Schultz. Chicago: University of Chicago Press.

Feldstein, M.S. 1966. A binary variable multiple regression method of analysing factors affecting perinatal mortality and other outcomes of pregnancy. J. Royal Stat. Soc. (Series A) 128:61-73.

Fuchs, V.R. 1974. Some Economic Aspects of Mortality in Developed Countries. In The Economics of Health and Medical Care, ed. M. Perlman. New York: Wiley.

Goldberger, A.S. 1964. Econometric Theory. New York: Wiley.

Gronau, R. 1974. The Effect of Children on the Housewife's Value of Time. In Economics of the Family, ed. T.W. Schultz. Chicago: University of Chicago Press.

Grossman, M. 1972. The Demand for Health: A Theoretical and Empirical Investigation. New York: Columbia University Press.

__________ 1974. The Correlation Between Health and Schooling. In Household Production and Consumption, ed. N.E. Terleckyj. New York: Columbia University Press.

Hulfin, M. and O'Hosson, M.O. 1971. Perinatal conditions of unwanted children. Acta Psychiat. Scand. 221:59-76.

Inman, R.P. 1976. The Family Provision of Childrens Health: An Economic Analysis. In The Role of Health Insurance in the Health Services Sector, ed. R.N. Rosett. New York: National Bureau of Economic Research.

Institute of Medicine, 1973. Infant Death: Analysis by Maternal Risk and Health Care. Washington, D.C.: National Academy of Sciences.

Johnston, J. 1963. Econometric Methods. New York: McGraw-Hill.

__________ 1972. Econometric Methods, 2nd ed. New York: McGraw-Hill.

Liebowitz, A. 1974. Education and the Allocation of Women's Time. In Education, Income, and Human Behavior, ed. F.T. Juster. New York: McGraw-Hill.

Michael, R.T. 1972. The Effect of Education on Efficiency in Consumption. New York: National Bureau of Economic Research.

__________ 1973. Education and the derived demand for children. J.P.E. Supplement (March/April).

Maurer, K.M. 1973. Dichotomous LOGIT Program. Unpublished computer program.

Morris, M.M.; Udry, R.J. and Chase, C.L. 1973. Reduction of low birth weight rates by the prevention of unwanted pregnancies. Amer. J. Public Health 63:935-938.

Muller, C.; Krasner, M.; and Jaffe, F.S. 1975. An index of insurance adequacy for fertility-related health care. Medical Care XIII: 25-36.

Nerlove, M. and Press, S.J. 1973. Univariate and Multivariate Log-Linear and Logistic Models. R-1306-EDA/NIH. Santa Monica: Rand Corporation.

Nortman, D. 1974. Parental Age as a Factor in Pregnancy Outcome and Child Development. Reports on Population/Family Planning no. 16.

O'Hara, D.J. 1972. Change in Mortality Levels and Family Decisions Regarding Children. R-914-RF. Santa Monica: Rand Corporation.

Pakter, J. and Nelson, F. 1974. Factors in the unprecedented decline in infant mortality in New York City. Bull. N.Y. Acad. Med. 50:839-868.

Quigley, J.M. 1976. Housing demand in the short run: an analysis of polytomous choice. Explorations in Economic Research 3:76-102.

Ross, S. 1974. The Effect of Economic Variables on the Timing and Spacing of Births. Ph.D. dissertation, Columbia University.

Shah, F.K. and Abbey, H. 1971. Effects of some factors on neonatal and postneonatal mortality. Milbank Mem. Fund Quar. XLIX:33-58.

Shapiro, S.; Schlesinger, E.R.; and Nesbitt, R.E.L., Jr. 1968. Infant, Perinatal, Maternal and Childhood Mortality in the United States. Cambridge: Harvard University Press.

Wegman, M.E. 1975. Annual summary of vital statistics, 1974. Pediatrics 56:960-6.

Welch, F. 1974. Sex of Children: Prior Uncertainty and Subsequent Behavior. R-1510-RF. Santa Monica: Rand Corporation.

Williams, R.L. 1974. Outcome-Based Measurements of Medical Care Output: The Case of Maternal and Infant Health. Ph.D. dissergation, University of California, Santa Barbara.

Williams, A.D. 1975. Fertility and Reproductive Loss. Ph.D. dissertation, University of Chicago.

Willis, R.J. 1972. The Economic Determinants of Fertility Behavior. Ph.D. dissertation, University of Washington.

__________ 1973. Economic theory of fertility behavior. J.P.E. Supplement (March/April).

Wray, J.D. 1971. Population Pressure on Families: Family Size and Child Spacing. Population/Family Planning, no. 9.