PAKISTAN INSTITUTE OF DEVELOPMENT ECONOMICS

Seminar Paper No. 19
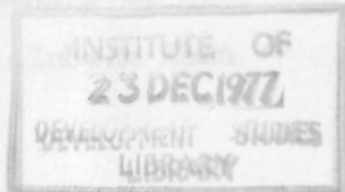
July 1977

ON THE CARE AND HANDLING OF REGRESSION SPECIFICATIONS
IN FERTILITY RESEARCH

Dennis De Tray

Zubeda Khan

ON THE CARE AND HANDLING OF REGRESSION SPECIFICATIONS
IN FERTILITY RESEARCH

Dennis De Tray

Zubeda Khan

June, 1977

Pakistan Institute of Development Economics
Islamabad

On the Care and Handling of Regression Specifications
in Fertility Research

Economists, sociologists and demographers must often attempt
to answer important questions with data not well suited to the
problem at hand. One example that crops up frequently in socio-
economic-demographic literature is the use of samples of women
whose ages span the entire fecund period to study the effects of
couples characteristics on "completed" fertility, or the demand
for children. In this case, the usual procedure is to control
either for age or duration of marriage, and to assume that timing
and spacing considerations can be ignored. Under this assumption
the stock of children at any point in time (any age), or differen-
ces in the level of that stock among families bears a one-to-one
correspondence to differences in completed fertility observed at
the end of the fertile period. In this note, we explore some of
the pitfalls that researchers may encounter when using such data,
especially in the absence of exact knowledge on the functional
form of the relationship between age or duration of marriage and
other variables thought to affect actual fertility.

We begin with a brief review of some regression results
published in an earlier issue of this journal, and then consider
several extensions of those results that on the surface appear
to produce interesting and policy-relevant new information about
factors that influence the number of children that couples have.
We then dissect these results with an eye toward assessing their
validity as measures of behavioural tendencies versus and alter-
native explanation based solely on a simple, though easily

overlooked, series of specification errors. Our analysis strongly suggest that the latter explanation -- specification error -- may account for many of the interesting findings extracted from our preliminary research, and offers a telling example of the care with which statistical analyses of individual and family behaviour must proceed.

Past Work

In a recently-published study, Afzal, Khan and Chaudry (AKC, hereeafter) analyse the fertility behaviour of a sample of Lahore women whose ages range from 17 to 70. Although their primary concern is with the relationship between (completed) fertility, age of marriage, and infant and child mortality the authors also consider the effect on observed fertility of a number of "socio-economic" factors such as wife's education and family income. The rationale for including these variables in an analysis of fertility can be drawn from a number of arguments based both on supply considerations like contraceptive knowledge, and on demand factors having to do with the costs and benefits associated with children. In what follows, we emphasize the demand side of these arguments, but the point of the paper is, in a sense independent of the school of thought to which the reader adheres.

Proponents of the deamdn interpretation of socio-economic variables state that there should exist a negative relationship between those factors that are associated with increasing costs of having and rearing children and the number of children couples have (other things equal, of course). If female schooling can

serve as a proxy for the value of a wife's time in activities other than child rearing, then children will be more expensive in households where the wife is highly educated than in households in which the wife has little or no education.[1] Thus, from demand considerations we predict a negative relationship between fertility and wife's schooling level.

The relationship between income and fertility is a much-debated one, but at its simplest level, if children are like most other items consumed by couples, as incomes rise, so will observed fertility. The controversy surrounding the relationship between income and fertility tends to focus on the conflict between this relatively straightforward theoritical prediction and the empirical fact that within most countries and mong most countries, rises in income are associated with declines in fertility. There are a number of explantions for this anomoly[2] but one possibility is that the gross negative correlation between income and fertility is not a casual phenomenon at all, but due to the positive correlation between income and the cost of rearing children; for example, the value of the mother's time in activities other than child rearing, and family income are likely to be positively corelated. Under this explanation, controlling for the price of children ought to change the relationship between income and fertility from negative to positive.

---

[1] For a more detailed discussion of the relationship between female schooling and fertility, see /¯5_7; the topic is covered in a somewhat more rigorous fashion in /¯ 4 _/.

[2] For a sophisticated theoretical argument as to why the partial relationship between income and fertility could be negative even if children were normal goods in the strict definition of this term, see /¯3_7.

We have gone briefly into some of the "theory" underlying our interpretation of the socio-economic variables available in the AKC study in order to give some indication of why we were initially dissatisfied with the regression results presented in $\underline{/\underline{}/}$. The AKC regression of interest has as its dependent variable children ever born, (CEB) and as independent variables duration of marriage (DM) age of marriage (AM) education of the father (EF) and the mother (EM), family income,[1] and the family's infant and child mortality rate (M--the ratio of infant and child deaths to children ever born). As the following equation indicates, of these variables, only duration of marriage and female education were significantly correlated with fertility:

$$CEB = 3.308+0.131 \; DM-0.234 \; EM-0.037 \; EF+0.486 \; M$$
$$(16.8) \qquad (3.7) \qquad (0.9) \qquad (1.19)$$
$$-0.020 \; AM$$
$$(1.0)$$
$$R^2 =0.373; \; F=79.5 \; \text{(t-ratio's in parentheses)}$$

$\underline{/}$Source: ? 2 $\underline{/}$ p. 203$\underline{/}$

The Extended Analysis

We began our analysis of the Lahore data by asking two questions which had not been considered in the two previous studies based on these data $\underline{/1/}$ and $\underline{/2/}$. The first question was whether the poor showing of the socioeconomic variables might not be a consequence of misspecification of the functional relationship between some of the explanatory variables and children ever born.

---

[1]Family income was actually dropped from the regressions reported in $/$ 2$/$ because the significance of its coefficient fell below a predetermined minimum level.

The second question was a natural consequense of the follow-
ing characteristic of the Lahore sample: of the 674 observations
in the working sample, 159 were drawn from Model Town, a rela-
tively well-to-do area of Lahore, while the remaining 515 obser-
vations were selected from Lahore Township, an area of predominantly
low income families. Thus, the second aim of our study was to
determine if the relationship between fertility and various socio-
economic characteristics of the family differ for families living
in high income areas as compared to families living in low income
districts.

To explore this possibility, the Lahore sample was divided
into the relevant subsamples, and separate regressions calculated
for each subsample. In addition, a number of alternative func-
tional forms were tested to determine whether the statistically
weak results reported in AKC were partly due to specification
error. Variables used in this analysis are drawn from the earlier
AKC work with some exceptions discussed below. Variable means
and standard deviations for each of the subsamples appears in
Table 1[1]

As is evident from table 1, the two subsamples differ sub-
stantially in their average socio-economic characteristics. The
average monthly income for Model Town is nearly five times that
of Lahore Township, with husband's and wife's average schooling
levels telling the same story. In contrast, the average numbers
of children ever born for the two communities are almost identical,

---

[1] Detailed descriptions of both the sample and the variables
can be found in either / 1 / or / 2 /

which is somewhat surprising given the substantial differences
other characteristics between the two areas. But in fact, fa:
size does differ. The similarity between mean numbers of chil
ren ever born for the two samples is an artifact of the diffe:
age distributions for women in Lahore Township and in Model T

TABLE-1
SUMMARY STATISTICS

Variable Name:	Model Town		Lahore Towns	
	Mean	s.d.	Mean	s.
Children ever born.	5.24	2.5	5.26	2.
Family Income	1073	1003	217	20
Wife's Education:				
None	0.26	0	0.91	
Less than Matric	0.21		0.05	
Matric or More	0.53		0.04	
Husband's Education:				
None	.006		0.64	
Less than Matric	.019		0.16	
Matric but less than BA	.371		0.15	
B.A.plus.	.604		0.05	
Age of wife	43.7	12.7	34.0	9.
Duration of Marriage	25.5	13.3	16.5	9.
Infant and child Mortality	0.11	0.17	0.19	0.
Sample size	159		515	

Lahore Township women are on the order of ten years younger th.
women in Model Town. Thus, in Model Town, women are, on avera;
at an age where their child bearing is virtually complete, in
contrast, women in the Lahore Township sample are likely to ha

borne several more children in the years following the survey. As will be seen, this age difference between the two samples plays an important role in explaining observed differences in "Behavioural" regression coefficients.

The hypotheses that guided our work on functional forms concern the nature of the relationship between family income and fertility, and between a couple's own infant and child mortality experiences and subsequent fertility. With regard to income, we considered the possibility that the effect of income on a couple's desired, and thus actual, fertility may depend in part on the level of income at which the couple lives. The relationships illustrated in Fig. 1 are two possible alternatives to the linear specification used in AKC and suggest that income has a positive influence on fertility at low income levels, but that at higher income levels the relationship either disappears, (line A) or becomes negative (line B).

Fig.1: Children ever born and Income: Alternative Relationships.
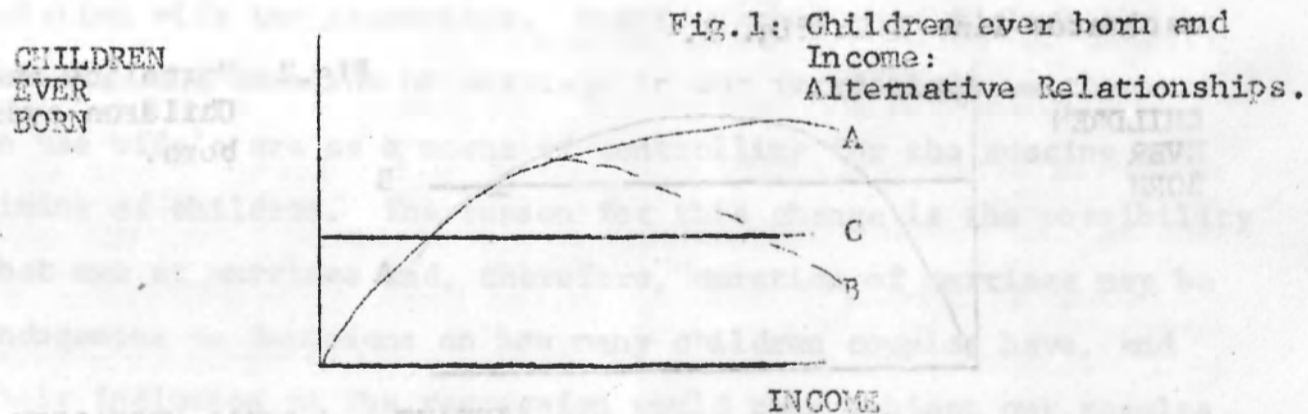
If line B represented the "true" relationship between income and fertility, then an estimate of the effect of income on fertility based on a linear income specification might well produce insignificant results (line C), even though there is in fact, a strong casual link between income and fertility.

The statistically weak showing of the mortality variab

the AKC analysis is, in many respects, much more puzzling tha

the lack of relationship between income and fertility. A str

positive correlation between fertility, as measured by childr

ever born, and infant mortality is one of the most consistant

observed relationships in studies of factors that influence f

size (see [ ] for a survey of this literature). Again, we

our hypotheses in terms of potential specification error -- c

the insignificant mortality coefficient be explained by the p

bility that, like income, the effect of mortality on fertili

is not linear (proportional), but differs by levels of mort.

If the true relationship between mortality and children ever l

can be represented by . line A in the following figure, then

AKC results may, again, be explained by the fact that mortalit

enters their equation in a linear form -- that is, their equat

estimates line B in fig. 2.

CHILDREN
EVER
BORN

Fig.2: Mortality an
Children eve
born.

INFANT and CHILD MORTALITY
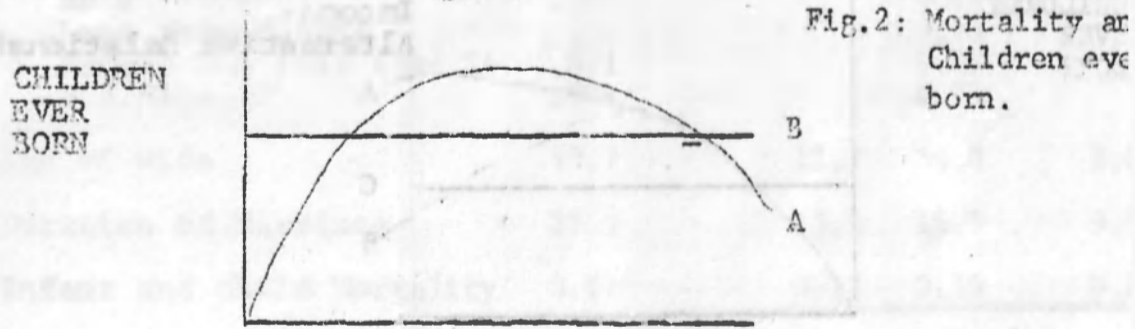
Although not essential to the purpose of this paper,

it is wroth noting that a reasonable case can be made to just

a relationship between fertility and mortality of the form giv

by line A in fig. 2. The behavioural link between children

ever born and mortality is often described in terms of a desir

the part of parents to replace lost children. But replacing child-
ren is far from a costless activity. Further, the cost and bene-
fit of replacing a lost child will depend in part on the proba-
bility of the subsequent birth surviving. If some
parents are, for reasons of health or environment, especially
prone to high levels of infant mortality, and if they base their
(perhaps implicit) estimate of the probability of a subsequent
birth surviving on their previous infant mortality experience,
then at high levels of infant mortality, parents may simply stop
trying to replace lost children. If mortality levels are suffi-
ciently high, these "discouraged parents" may also revise down-
ward their desires for living children, which could account for
the negative portion of line A in fig. 2.

Age, Income and Fertility

In exploring the possibility of a nonlinear relationship
between income and fertility, we followed the previous AKC for-
mulation with two exceptions. First, rather than include age at
marriage and duration of marriage in our regressions we choose
to use wife's age as a means of controlling for the spacing and
timing of children. The reason for this change is the possibility
that age at marriage and, therefore, duration of marriage may be
endogenous to decisions on how many children couples have, and
their inclusion in the regression would thus subject our results
to simultaneity bias.[1] Second, husband's and wife's education
levels are entered as a series of categorical variables, a minor
deviation from the earlier work on these data. And finally, we

---

[1] See [57] for a discussion of the casual relationship age of
marriage and fertility.

concentrated our initial efforts on the larger of the two samples, that for Lahore Township, and will report results for the Model Town sample only when these results add to the discussion.

Although there is seldom strong theoretical justification favouring one functional form over another, common sense leed us to explore first functional forms which approximate line A in Fig. 1, that is, which produce a positive but declining relationship between income and fertility. A frequently used candidate in this class of functions is a semi-log transformation where the natural log of income is treated as an explanatory variable.

The result of our initial test of this functional form is given in eq. 1 of table 2, and it seems that income does indeed affect couple's fertility behaviour. The coefficient of the natural log of family income is positive and significant at the five percent level, indicating that the lower the income level, the greater the effect of income on fertility.

For the Model Town sample, the equivalent regression produced an essentially zero coefficient on the log of income which we took as further confirmation of our hypothesis that the effect of income on fertility declines as the level of income rises; that is , the mean level of income in Model Town is sufficiently high so that most couples in that sample fall in the flat part of the income-fertility curve.

A relatively high income elasticity of fertility among the poor has important and somewhat discouraging policy implications for a country attempting to slow population growth rates. To add to the confidence with which our results could be viewed, we

attemped to assess the robustness of our finding by estimating
several alternative specifications of eq.1, and among these was
the regression given in eq. 2 of table 2. The justification for
including the square of wife's age as an explanatory variable is
straightward -- the relationship between age and fertility must
decline as women reach the end of their fecund period -- but the
important result in this new specification concerns its effect
on the coefficient of the log of family income. With the addition
of wife's age squared, the income coefficient falls in absolute
value, and its t-ratio drops to well below conventional significance
levels

Whether this new finding represents a rejection of our
income hypothesis depends on why the change in the income coeffi-
cient occurred. One explanation could be that we have pushed these
data too hard, and that a larger sample, or a better measure of
income would again produce a significant income effect. A more
concrete explanation, however, can be drawn from the age distri-
bution of women in the Lahore Township data in combination with a
subtle form of omitted-variable bias.

To understand why the income results in eq. 1 could be
the result of ommitted-variable bias, it is first necessary
to review the relationship between income and age. There is both
theoretical and much empirical evidence to suggest that over a
family's life cycle, earnings and therefore income rise but at
an ever-declining rate.[1]

---

[1] See /6/ for a study based on Pakistani data.

TABLE-2

REGRESSIONS RESULTS: LAHORE TOWNSHIP FORM I
Dependent Variable: **Children ever Born.**

EXPLANATORY VARIABLES[a]	Eq. 1.	Eq. 2.
Wife's Age	0.16 (13.6)	0.61 (9.6)
(Wife's Age)^2		-0.0059 (7.2)
Wife's Education: Less than Matric (EDW1)	-1.29 (2.5)	-1.20 (2.4)
Matric Plus (EDW2)	-1.23 (1.8)	-0.90 (1.4)
Husband's Education: Less than Matric	-0.07 (0.2)	-0.13 (0.5)
Matric but less than B.A.	0.12 (0.2)	0.29 (0.9)
B.A. Plus	0.86 (0.2)	0.26 (0.4)
Ln (Family Income)	0.56 (1.9)	0.35 (1.2)
Mortality	0.75 (1.5)	0.92 (2.0)
Intecent	3.23	-10.03
R^2	0.33	0.39
F	31.2	36.2
N	515	-10.03

[a] t-ratios in parentheses.

One such hypothetical age-income profile is plotted in fig. 3.

Income

Fig. 3 Hypothetical Age-Income
Profile.

Age
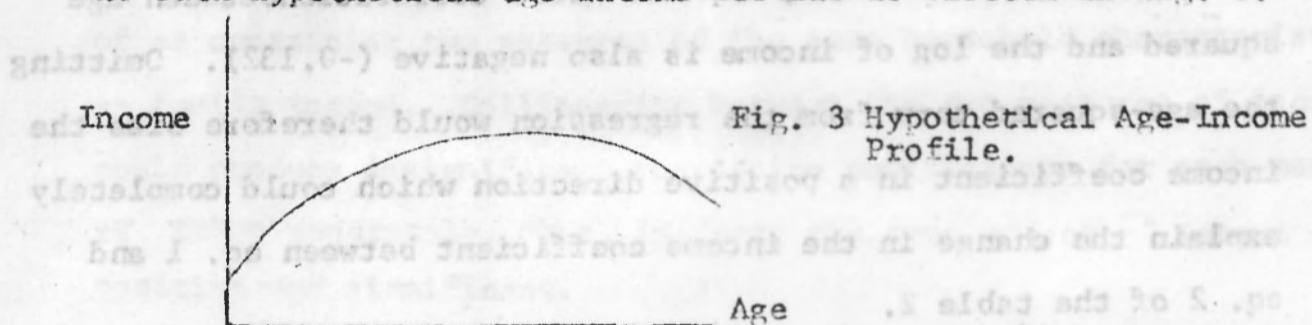
Now, we have already established that the relationship between
wife's age and fertility is also likely to be positive but declining.
Thus, if we were to leave age out of our regression specification
entirely, we would run the risk of having the very strong age-
fertility relationship picked up by income or some transformation
of income. That is, with age excluded, we could estimate a
significant and positive coefficient for income that was due
entirely to the correlation between income and wife's age, and
not at all to the behavioural influence of income on fertility.

In the specification for eq. 1, wife's age has been included
as an explanatory variable, but not in the appropriate functional
form. In a sense, we have ommitted the variable "wife's age
squared" and it is the correlation of the log of family income with
this ommited variable that could account for the positive and
significant income coefficient in eq. 1. The direction of the
bias caused by an omitted variable depends on the correlation
between the omitted variable and the included variables, and on
the partial correlation between the omitted variable and the
dependent variable in the regression. In this case, both the
predicted and estimated partial correlation between age squared

and fertility is negative, and the cross correlationbetween ag
squared and the log of income is also negative (-0.132). Omit
the age-squared term from the regression would therefore bias
income coefficient in a positive direction which could complet
explain the change in the income coefficient between eq. 1 and
eq. 2 of the table 2.

Having arrived at at least one explantion for the change
the income coefficient between eqs 1 and 2, we now consider the
issue of why income is not significantly related to fertility v
in theory it ought to be. As we stated above, measurement erro
tends to bias coefficients towards zero, and family income is l
to be very poorly measured. But, again we have no way of testin
this hypothesis short of doing additional field work. However,
is one explanation for income's poor showing that can be easily
tested using these data.

The regression specification used in AKC and by us conta
measures of both husband's and wife's education levels. In our
earlier discussion we considered the role of female education a
a proxy for the value of the wife's time, and therefore the "co:
of rearing children, but we offered no justification for the in
sion of male schooling in the regression. While there are a nu
of possible reasons for treating husband's schooling as an expl:
natory variable in a fertility regression, the most common just:
fication revolves around the correlation between schooling and
income. When better measures of income are not available, husb:
education is often used as a proxy for the family's wealth statt

or income level. But in the Lahore data, we have a direct measure
of income; our regressions to this point, therefore, can be thought
of as containing two measures of the same household characteristic
-- family income. Collinearity between the two measures of income
could produce insignificant regression coefficients for each measu-
re taken separately, when, in fact, the true effect of income is
positive and significant.

Table 3 contains the results of our test of this hypothesis
for several different specifications of the wife's age variable.
The log income coefficient is significant at conventional levels
in three of the four specifications, suggesting that collinearity
may, indeed, have been a problem in the earlier specifications.

These new results raise an additional question: which of
the specifications for wife's age produces the "best" results in
the sense of a least-biased estimate of the effect of income on
fertility? Based on the criteria of t-ratios and F statistics,
of the four specifications in table 3, the log specifications
seems to fit best. However, since we cannot choose among the non-
linear age specifications on a priori grounds, we decided that
the least costly specification in terms of bias ought to be one
in which age is allowed to seek its own functional relationship
with the dependent variable. To do this, we divided the wife's
age variable into six short linear segments defined so that they
would produce a continuous if disjointed profile of the relation-
ship between age and children ever born.[1]

[1.]The linear segments were defined as follows: AGE1=actual age
for those women between 17 and 21, or 21 for all women over
21; AGE2=0 for women aged 21 or less, 1 through 5 for women
aged 22 through 26 respectively, and 5 for all women over 26;
AGE3=0 for all women aged 26 or less, 1 through 7 for women
aged 27 through 33 respectively, and 7 for all women over 33;
and so on for age segments 34 through 39, 40 through 45 and 46 plus.

The results of this new specification are given in the follow equation:

$$CEB = -11.3 + 0.56 \; AGE1 + 0.16 \; AGE2 + 0.32 \; AGE3 + 0.22 \; AGE4$$
$$(1.6) \qquad (1.2) \qquad (5.1) \qquad (2.5)$$

$$+0.07 \; AGE5 \; -0.02 \; AGE6 \; -1.11 \; EDW1 \; -0.79 \; EDW2$$
$$(0.9) \qquad (0.6) \qquad (2.3) \qquad (1.3)$$

$$+0.44 \; LN \; (INCOME) + 0.78 \; MORTALITY$$
$$(1.7) \qquad\qquad (1.7)$$

$$R^2 = 0.40: \; F = 33.0 \quad N = 515$$

where AGE variables are defined in footnote 1 and EDW variabl table 3. As Fig. 4, and a quick comparison of this equation the equations given in table 3 indicate, the -nconstrained re ship between wife's age and fertility tends to closely approx a quadratic, and, thus, the effect of income on children ever born while positive, is not significant at conventional level.

Fig. 4 Age-Fertility Pro
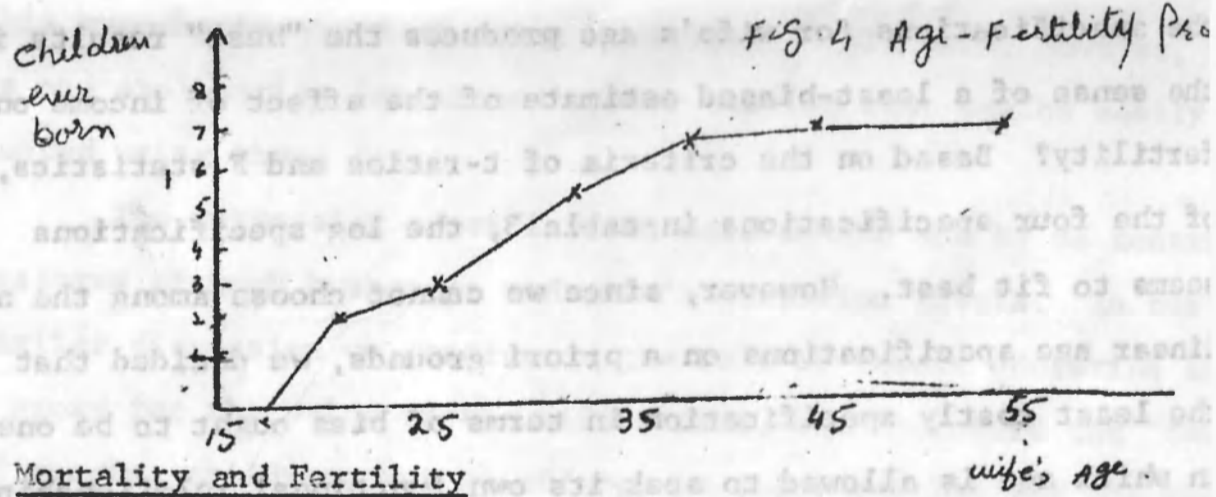
Mortality and Fertility

As table 3 indicates the statistical strength of the inf and child mortality coefficient depends in part on the form ir which wife's age enters the regression. In general, the effec of mortality on children ever born is not significant at conve tional levels except when wife's age enters the regression as quadratic. To test our hypothesis that mortality effects fert

TABLE-3

CHILDREN EVER BORN REGRESSION: FORM II

Explanatory Variables [a]	Equation Number			
	1.	2.	3	4.
Wife's Age	0.16 (13.9)	0.60 (9.6)		
(Wife Age)^2		-0.0058 (7.1)		
Ln (Age)			6.03 (15.2)	
$\frac{\Lambda}{A}$				2.0 (14.6)
Wife's Education				
Less than Matric (EDW1)	-1.27 (2.6)	-1.10 (2.3)	-1.2 (2.4)	-1.2 (2.5)
Matric plus (EDW2)	-1.16 (1.9)	-0.72 (1.2)	-0.94 (1.6)	-1.04 (1.7)
Ln (Income)	0.6 (2.1)	0.43 (1.6)	0.53 (2.0)	0.56 (2.0)
Mortality	0.73 (1.5)	0.88 (1.9)	0.71 (1.5)	0.71 (1.5)
Intercept	-3.36	-10.28	-18.57	-9.32
R^2	0.33	0.39	0.37	0.35
F	50.1	54.3	58.4	54.8
N	515	515	515	515

[a] t-ratios in parentheses.

in some nonlinear fashion, and that this accounts for the frequent insignificance of the mortality coefficient, we took the square of the infant mortality variable and entered it as an additional explanatory variable. The results of this regression are given as eq. 2 of table 4 (eq. 1 in that table is presented for comparative purposes only).

By all conventional measures, we appear to have discovered a very strong quadratic relationship between a couple's infant mortality experience and the number of children born by the wife. The addition of the mortality-squared variable significantly raises the proportion of variance explained by the regression, and the coefficients on the two mortality variables are highly significant, both having t-ratios inexcess of seven. This relationship has been graphed in fig. 5 using mean or modal values of other variables in the regression.

Fig. 5 Children ever born and mortality
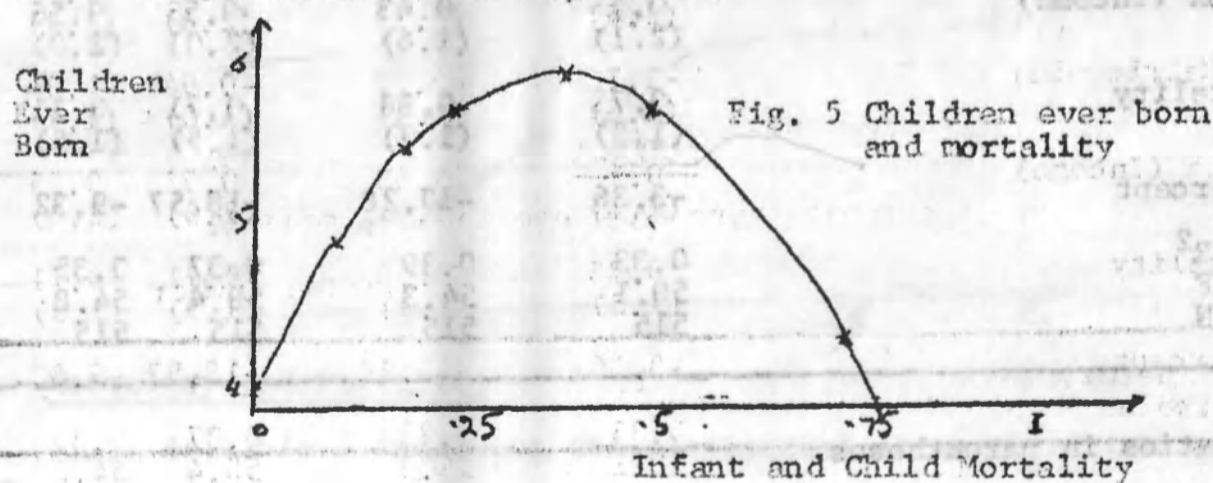
Without going into details it seems fairly clear that were fig.5 an illustration of a behavioural phenomenon, our findings would have important implications for the future growth of population in Pakistan, and for the distribution of that growth among different types of couples in the population.

TABLE-4

FERTILITY AND MORTALITY

Explanatory Variables:	Eq. No.	
	1.	2.
Wife's Age	0.60	0.54
	(9.6)	(9.1)
(Wife's Age)^2	-0.0058	-0.0052
	(7.1)	(6.7)
Wife's Education:		
Less than Matric	-1.10	-0.96
	(2.3)	(2.1)
Matric, plus	-0.72	-0.42
	(1.2)	(0.8)
Ln (Income)	0.43	0.37
	(1.62)	(1.47)
Mortality	0.88	8.1
	(1.9)	(7.7)
(Mortality)^2		-10.5
		(7.6)
Intercept	-3.36	-9.23
R^2	0.39	0.45
F	54.3	59.8
N	515	515

Because the results presented in table 4 are almost too g

that is, statistically too strong for mic data, we decided that

closer look at possible non-behavioural links between mortality

and fertility was in order. The most obvious explanation, tha

of multicollinearity between mortality and mortality squared, c

not be tested directly because the computer program used in our

analysis does not supply the necessary information on the variat

and covariance of the estimated coefficients, but, it seems to

us that the real answer lies elsewhere in any case.

Let us assume that there is no behavoural link between

mortality and fertility -- is there any reason why we might stil

except a relationship of the form illustrated in fig. 5? The

answer is yes, and the reason has to do with the probability of

observing families with different numbers of children ever born

at the extremes of the infant and child mortality spectrum. If

a family has only one child, that is one birth, then the probab

of observing that couple at the upper extreme of the infant mor-

tality variable (i.e. a value of one) is approximately 0.188,

which is the mean infant mortality rate for the sample as a whole

Similarly, the probability that that family has a zero value for

the mortality variable is 0.812 (1.0-.188). Now, consider the

other extreme of the family size distribution, a family with six

children. If mortality rates do not vary significantly by family

size, then probability that this second family will be at th

upper end of the mortality range is 0.000044 (0.188⁶) and at the

lower end, 0.29 (0.8126). Thus, without any behavioural feedback

between mortality and fertility, we can completely explain the shape of the curve illustrated in fig. 5. Both its curvilinear nature and the fact that the intercept at a zero infant mortality rate is higher than the intercept at an infant mortality rate of one can be explained solely on the basis of the relative probabilities of large and small families being observed at extreme values of the infant mortality variable.

The point of the preceding discussion is that while there are reasonable arguments to support the contention that the relationship between mortality and fertility is nonlinear, it would be virtually impossible to test that hypothesis with data like those used in this study. There are, of course, other more subtle methods of exploring the effect of infant and child mortality on fertility, for example by looking at birth intervals following an infant or child death, but directly assessing the nature of the link between <u>completed</u> fertility and mortality would appear to be an impossible task.[1] Further, unless one recognises the nature of the nonbehavioural relationship between mortality and fertility, one could be led to seriously biased policy conclusions regarding the effect on family behaviour of a decline in mortality.

----

[1] Note that this statement holds true not just for the nonlinear exploration described above, but for any estimate of the partial correlation between infant mortality and fertility. That part of the curve in fig. 6 falling between a zero infant mortality rate and, say, one standard deviation above the mean rate could produce a positive and significant linear coefficient since most of the sample observation will fall in that range.

Summary and Conclusions.

In this paper, we have tried to demonstrate some of tl
difficulties researchers face when analysing variations in cc
leted fertility using data on women who are still in their ch
bearing years. Although the purpose of the paper is mainly p
cautionary, at least one important implication for future res
can be drawn from our analysis. With regard to the effect of
in income on the total number of children a couple has, our r
are suggestive of a more complex relationship between fertili
income than has been estimated in many previous regression sp
fications. Further, although our "best" results (see p. 16) a
statistically too weak to serve as a basis for policy conclusi
they do point toward the possibility that a secular rise in in
may initially exacerbate the "population problem" in Pakistan
increasing fertility rates especially among low income familie

Our analysis of the relationship between infant and chil
mortality and fertility indicates that many previous studies o
this relationship may have produced seriously biased results, a
therefore, misleading policy implications. Although it may be
ssible to develop methodologies to overcome the problem raised
our mortality analysis, in our opinion, nothing short of comple
birth histories for women who have completed child bearing (and
the accompanying socio-economic information) would suffice for
an analysis of the behavioural links between fertility and mort

With the increased availability of micro data on individi
and families in Pakistan, there is likely to be a considerable

outpouring of regression analyses of the type outlined above. While
much can be learned about human behaviour from these and related
empirical exercises, it is obvious from our work that a great deal
of false information can also result. Thus, the principal recomm-
endation that can be drawn from this study has to do with methodo-
logy rather than substance -- in any analysis of individual behaviour,
each significant finding should receive a second and even a third
look if spurious results are to be kept to minimum.

REFERENCES

1. Afzal, M. Zubeda Khan and N.A. Chaudhry, "Age at Marriage, Fertility and Infant-Child Mortality in a Lahore Suburb (Part I)". Pakistan Development Review, Vol. XV, No.1, Spring, 1976

2. Afzal, M. Zubeda Khan and N.A. Chaudhry, "Age at Marriage, Fertility and Infant-Child Mortality in a Lahore Suburb (Part II)". Pakistan Development Review, Vol. XV, No.2, Summer, 1976.

3. Becker, Gary S. and Nigel Tomes. "Child Endowments and the Quantity and Quality of Children". Journal of Political Economics. Vol 84, No. 4. Part 2. August, 1976.

4. De Tray Dennis. "Population Growth and Educational Policies: An Economic Perspective". in Population and Development, Ronald Ridker ed. Baltimore, The Johns Hopkins University Press. 1976.

5. _____, "Age of Marriage a   ertility: A Policy Review". Pakistan Development Rev   Vol XVI, No. 1. Spring 1977.

6. Hamdani, Khalil. "Education and the Income Differential: An Estimation for Rawalpindi City". Pakistan Development Review. Vol. XVI, No 2. Summer, 1977.

7. Schultz, T. Paul. "Interrelationships between Mortality and Fertility" in Population and Development, Ronald Ridker, ed. Baltimore. The Johns Hopkins University Press. 1976.