On the Care and Handling of Regression Specifications in Fertility Research

DENNIS DE TRAY and ZUBEDA KHAN*

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 on the demand for children. In this case, the usual procedure is to control either for age or for duration of marriage, and to assume that issues concerned with timing and spacing of children can be ignored. Under this assumption, differences in the level of the stock of children among families at any point in time (any age) bear a one-to-one correspondence to differences in completed fertility observed at the end of the fertile period. In this paper, we explore some of the pitfalls researchers may encounter when using data with this characteristic, especially in the absence of exact knowledge of 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 couples have. We then dissect these results with an eye toward assessing their validity as measures of behavioural tendencies versus an alternative explanation based solely on a simple though easily overlooked, series of specification errors. Our analysis strongly suggests 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 some of the dangers associated with statistical analyses of individual and family behaviour.

---

*The authors are, respectively, Advisor, The Pakistan Institute of Development Economics and Associate, The Population Council; and Staff Demographer, The Pakistan Institute of Development Economics. They would like to thank Mohammad Afzal, Jeff Hammer, David Oot, Steven Sinding, and the participants of PIDE's seminar series for useful comments on an earlier draft, and Syed H.H. Naqavi for editorial and stylistic suggestions.

PAST WORK

In a recently-published study, Afzal, Khan and Chaudry (AKC, here-in after) 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 may adhere.

Proponents of the demand interpretation of the relationships between socio-economic variables and fertility state that there should exist a negative partial correlation between those factors that are associated with increasing costs of having and rearing children and the number of children couples have (other things being 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, normal goods, as incomes rise, so, too, will observed fertility. The controversy surrounding the relationship between income and fertility tends to focus on the conflict between this relatively straightforward theoretical prediction and the empirical fact that within most countries and among most countries, rises in income are associated with declines in fertility. There are a number of explanations for this anomaly[2] but one possibility is that the gross negative correlation between income and fertility is due not to a causal link between these variables, but 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 correlated. Under this explanation, holding constant differences in the price of children ought to change the relationship between income and fertility from negative to positive.

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 [2]. The AKC regression of interest has as its dependent variable children ever born (CEB) and as independent variables duration of marriage (DM), age at marriage (AM), education of the father (EF) and of the

[1]For a more detailed discussion of the relationship between female schooling and fertility, see [5]; 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].

mother (EM), family income[3] 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:

$$\begin{aligned} CEB = {} & 3.308 + \underset{(16.8)}{0.131DM} - \underset{(3.7)}{0.234EM} - \underset{(0.9)}{0.037\ EF} + \underset{(1.19)}{0.486\ M} \\ & - \underset{(1.0)}{0.020\ AM} \end{aligned}$$

$R^2 = 0.373$; $F = 79.5$ (t-ratios in parentheses)

Source: [2, p. 203]

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 [1, 2]. The first question was whether the poor showing of the socio-economic variables was a consequence of misspecification of the functional relationships between some of the explanatory variables and children ever born. The second question was a natural consequence of the following characteristic of the Lahore sample: of the 674 observations in the working sample, 159 were drawn from Model Town, a relatively well-to-do area of Lahore, while the remaining 515 observations were selected from Lahore Township, an area of predominantly low-income families. Thus, the second aim of our study was to explore the relationships between fertility and various socio-economic characteristics of the families for couples living in high-income areas as compared to those living in low-income districts.

To undertake this analysis the Lahore sample was divided into relevant subsamples, and separate regressions calculated for each subsample. In addition, a number of alternative functional 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 appear in Table 1.[4]

As is evident from Table 1, the two subsamples differ substantially in their average socio-economic characteristics. The average monthly income for Model Town is nearly five times that of Lahore Township, with husbands' and wifes' average schooling levels telling the same story. In contrast, the average numbers of children ever born for the two communities are almost identical, which is surprising given the substantial differences in other characteristics between the two areas. But in fact, family size does differ. A closer examination of the data indicates that the similarity between mean numbers of children ever born for the two samples is an artifact of the different age distributions for women in Lahore Township and in Model Town.

---

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

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

Table 1

Summary Statistics: Lahore Sample

Variable Name	Model Town		Lahore Township	
	Mean	s.d.	Mean	s.d.
Children ever born (number)	5.24	2.5	5.26	2.9
Family Income (rupees per month)	1073	1003	217	202
Wife's Education:				
None	0.26		0.91	
Less than Matric	0.21		0.05	
Matric or More	0.53		0.04	
Husband's Education:				
None	0.01		0.64	
Less than Matric	0.12		0.16	
Matric but less than B.A.	0.37		0.15	
B.A. plus.	0.60		0.05	
Age of Wife (Years)	43.7	12.7	34.0	9.6
Duration of Marriage (Years)	25.5	13.3	16.5	9.9
Infant and Child Mortality	0.11	0.17	0.19	0.23
Sample size	159		515	

Note : Variables not otherwise designated are proportions.

Lahore Township women are, on average, ten years younger than women in Model Town. Thus, in Model Town, women are, for the most part, at an age when their child bearing is virtually complete; in contrast, women in the Lahore Township sample are likely to have 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 relationships between family income and fertility, and between a couple's own infant and child mortality experience 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 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 as income rises, the relationship either disappears (line A) or becomes negative (line B).

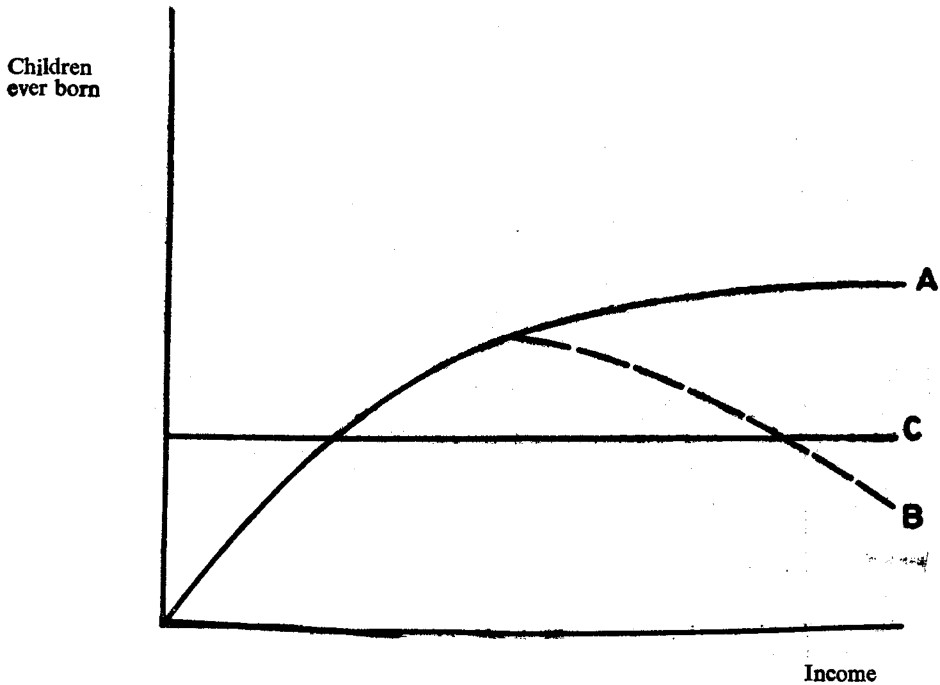

Fig. 1
Children ever born and Income
Alternative Relationships

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 causal link between income and fertility.

The statistically weak showing of the mortality variable in the AKC analysis is, in many respects, much more puzzling than the lack of relationship between income and fertility. A strong positive correlation between fertility, as measured by children ever born, and infant mortality is one of the most consistently-observed relationships in studies of factors that influence family size (see [7] for a survey of this literature). Again, we posed our hypothesis in terms of potential specification error—could the insignificant mortality coefficient be explained by the possibility that, like income, the effect of mortality on fertility is not linear (proportional) but differs by levels of mortality? If the true relationship between mortality and children ever born can be represented by line A in Fig. 2, then the AKC results may, again, be explained by the fact that mortality enters their equation in a linear form—that is, their equation estimates line B in Fig. 2.

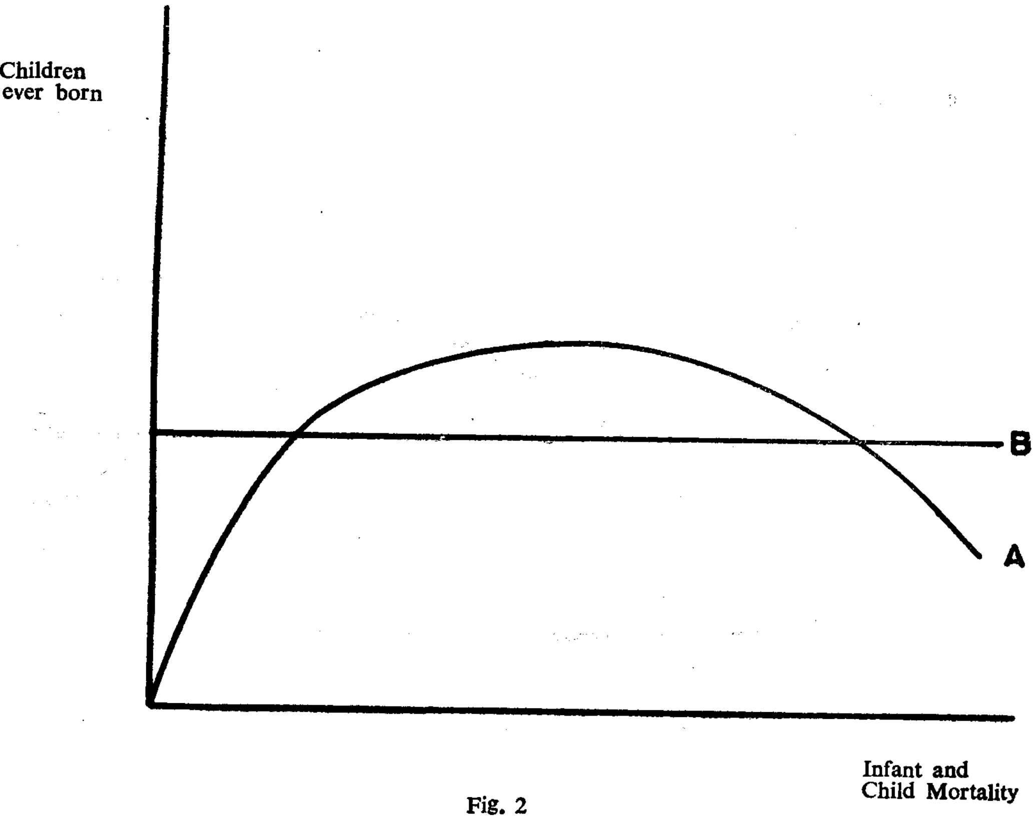

Fig. 2
Mortality and Children ever born

Although not essential to the purpose of this paper, it is worth noting that a reasonable case can be made to justify a relationship between fertility and mortality of the form given by line A in Fig. 2. The behavioural link between children ever born and mortality is often described in terms of a desire on the part of parents to replace lost children. But replacing children is far from a costless activity. Further, the cost and benefit of replacing a lost child will depend in part on the probability 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 sufficiently high, these "discouraged parents" may also revise downward 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 formulation with two exceptions. First, rather than include age at marriage and duration of marriage in our regressions, we 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 decisionis on the numbers of children couples want, and their inclusion in the regression would thus subject our results to simultaneity bias.[5] Secondly, 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 concentrated our initial efforts on the larger of the two samples—that for Lahore Township.

Although there is seldom strong theoretical justification for one functional form against another, common sense led us to explore first functional forms which approximate line A in Fig. 1, that is, those 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 in which 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 a 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. We took this as a 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 attempted 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 straightforward—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 the wife's age squared variable, the income coefficient falls in absolute value, and its t-ratio falls 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 coefficient 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.

---

[5]See [5] for a discussion of the causal relationship between age at marriage and fertility.

Table 2

Regressions Results, Lahore Township Form I
Dependent Variable: Children ever Born

Explanatory Variables	Equation 1	Equation 2
Wife's Age	0.16	0.61
	(13.6)	(9.6)
(Wife's Age)^2		−0.0059
		(7.2)
Wife's Education:		
Less than Matric (EDW1)	−1.29	−1.20
	(2.5)	(2.4)
Matric, Plus (EDW2)	−1.23	−0.90
	(1.6)	(1.4)
Husband's Education:		
Less than Matric	−0.07	−0.13
	(0.2)	(0.5)
Matric but less than B.A.	0.12	0.29
	(0.2)	(0.9)
B.A. Plus	0.06	0.26
	(0.2)	(0.4)
Ln (Family Income)	0.56	0.35
	(1.9)	(1.2)
Mortality	0.75	0.92
	(1.5)	(2.0)
Intercept	3.23	−10.03
R^2	0.33	0.39
F	31.2	36.2
N	515	515

Note: t-ratios are given in parentheses.

A more concrete explanation, however, can be drawn from the age distribution 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 omitted-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 the course of a family's life cycle, earnings and therefore income rise but at an ever-declining rate.[6] One such hypothetical age-income profile is plotted in Fig. 3.

[6]See [6] for a study based on Pakistan data.

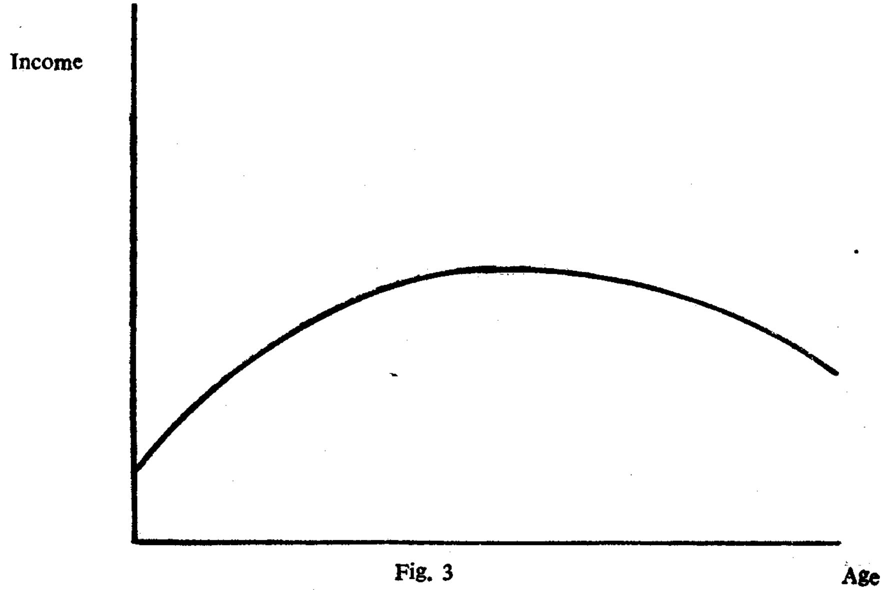

Fig. 3
Hypothetical Age-Income Profile

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 income or some transformation of income pick up the very strong age-fertility relationship. 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 omitted the variable "wife's age squared" and it is the correlation of the log of family income with this omitted 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 correlations between age square and fertility are negative, and the gross correlation between age squared and the log of income is also negative (—0.132). Omitting the age-squared term from the regression would therefore bias the income coefficient in a positive direction which could completely explain the change in the income coefficient between Eq. 1 and Eq. 2 of the Table 2.

Having arrived at one explanation for the change in the income coefficient between Eqs. 1 and 2, we now consider the issue of why income is not significantly related to fertility when in theory it ought to be. As we stated above, measurement error tends to bias coefficients towards zero, and family income is likely to be very poorly measured. But, again we have no way of testing this hypothesis short of doing additional field work. However, there 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 contains measures of both husband's and wife's education levels. In our earlier discussion we considered female education as a proxy for the value of the wife's time, and therefore for the "cost" of rearing children, but we offered no justification for the inclusion of male schooling in the regression. While there are a number of possible reasons for treating husband's schooling as an explanatory variable in a fertility regression, the most common justification revolves around the correlation between schooling and income. When better measures of income are not available, husband's education is often used as a proxy for the family's wealth status or income level. But, in the Lahore data, we have a direct measure of income; thus, our regressions, to this point, 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 measure 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 equations.

These new results raise an additional question: which of the specifications for wifes 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 specification seems to fit best. However, since we cannot choose among the nonlinear 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 so defined as to produce a continuous if disjointed profile of the relationship between age and children ever born.[7] The results of this new specification are given in the following equation:

$$\begin{aligned} CEB = {} & 11.3 + \underset{(1.6)}{0.56}\ AGE1 + \underset{(1.2)}{0.16}\ AGE2 + \underset{(5.1)}{0.32}\ AGE3 + \underset{(2.5)}{0.22}\ AGE4 \\ & + \underset{(0.9)}{0.07}\ AGE5 - \underset{(0.6)}{0.02}\ AGE6 - \underset{(2.3)}{1.11}\ EDW1 - \underset{(1.3)}{0.79}\ EDW2 \\ & + \underset{(1.7)}{0.44}\ LN\ (INCOME) + \underset{(1.7)}{0.78}\ MORTALITY \end{aligned}$$

$R^2 = 0.40$: $F = 33.0$ $N = 515$

[7]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 age 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.

AGE variables are as defined in footnote 1 and EDW variables are as defined in Table 3. As Fig. 4 and a quick comparison of this equation with the equations given in Table 3 indicate, the unconstrained relationship between wife's age and fertility tends to approximate a quadratic closely and, thus, the effect of income on children ever born while positive, is not significant at conventional levels.

Table 3

Children Ever Born Regressions: Form II

Explanatory Variables	Equation 1	2	3	4
Wife's Age	0.16	0.60		
	(13.9)	(9.6)		
(Wife's Age)^2		−0.0058		
		(7.1)		
Ln (Age)			6.03	
			(15.2)	
$\sqrt{\text{Age}}$				2.0
				(14.6)
Wife's Education				
Less than Matric (EDW1)	−1.27	−1.10	−1.2	−1.2
	(2.6)	(2.3)	(2.4)	(2.5)
Matric, plus (EDW2)	−1.16	−0.72	−0.94	−1.04
	(1.9)	(1.2)	(1.6)	(1.7)
Ln (Income)	0.60	0.43	0.53	0.56
	(2.1)	(1.6)	(2.0)	(2.0)
Mortality	0.73	0.88	0.71	0.71
	(1.5)	(1.9)	(1.5)	(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

Note: t-ratios are in parentheses.

Mortality and Fertility

As Table 3 indicates, the statistical strength of the infant and child mortality coefficient depends in part on the form in which the wife's age variable enters the regression. In general, the effect of mortality on children ever born is not significant at conventional levels except when the wife's age variable enters the regression as a quadratic. To test our hypothesis that mortality affects fertility 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.)

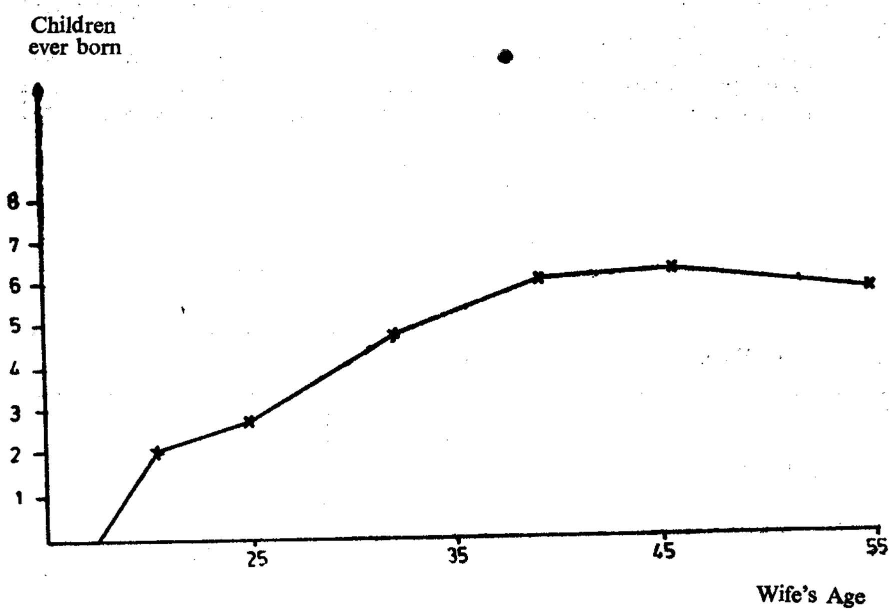

Fig. 4

Age-Fertility Profile

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

Without going into detail, 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.

Because the results presented in Table 4 are almost too good, that is, statistically too strong for micro data, we decided that a closer look at possible nonbehavioural links between mortality and fertility was in order. The most obvious explanation—that of multicollinearity between mortality and mortality squared—could not be tested directly because the computer programme used in our analysis does not supply the necessary information on variances and covariances of estimated coefficients, but it seems to us that the real answer lies elsewhere in any case.

Table 4

Fertility and Mortality: Alternative Functional Forms

Explanatory Variables	Equation 1	Equation 2
Wife's Age	0.60 (9.6)	0.54 (9.1)
(Wife's Age)^2	−0.0058 (7.1)	−0.0052 (6.7)
Wife's Education:		
Less than Matric	−1.10 (2.3)	−0.96 (2.1)
Matric, plus	−0.72 (1.2)	−0.42 (0.8)
Ln (Income)	0.43 (1.62)	0.37 (1.47)
Mortality	0.88 (1.9)	8.1 (7.7)
(Mortality)^2		−10.5 (7.6)
Intercept	−10.38	−9.23
R^2	0.39	0.45
F	54.3	59.8
N	515	515

Let us assume that there is no *behavioural* link between mortality and fertility. Is there any reason why we might still expect 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 couple has only one child, that is one birth, then the probability of observing that couple at the upper extreme of the infant mortality-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 the couple has a zero value for the mortality variable is 0.812 (1.0—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 the probability that this second family will be at the upper end of the mortality range is $(0.188)^6 = 0.000044$ and at the lower end, $(0.812)^6 = 0.29$. 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

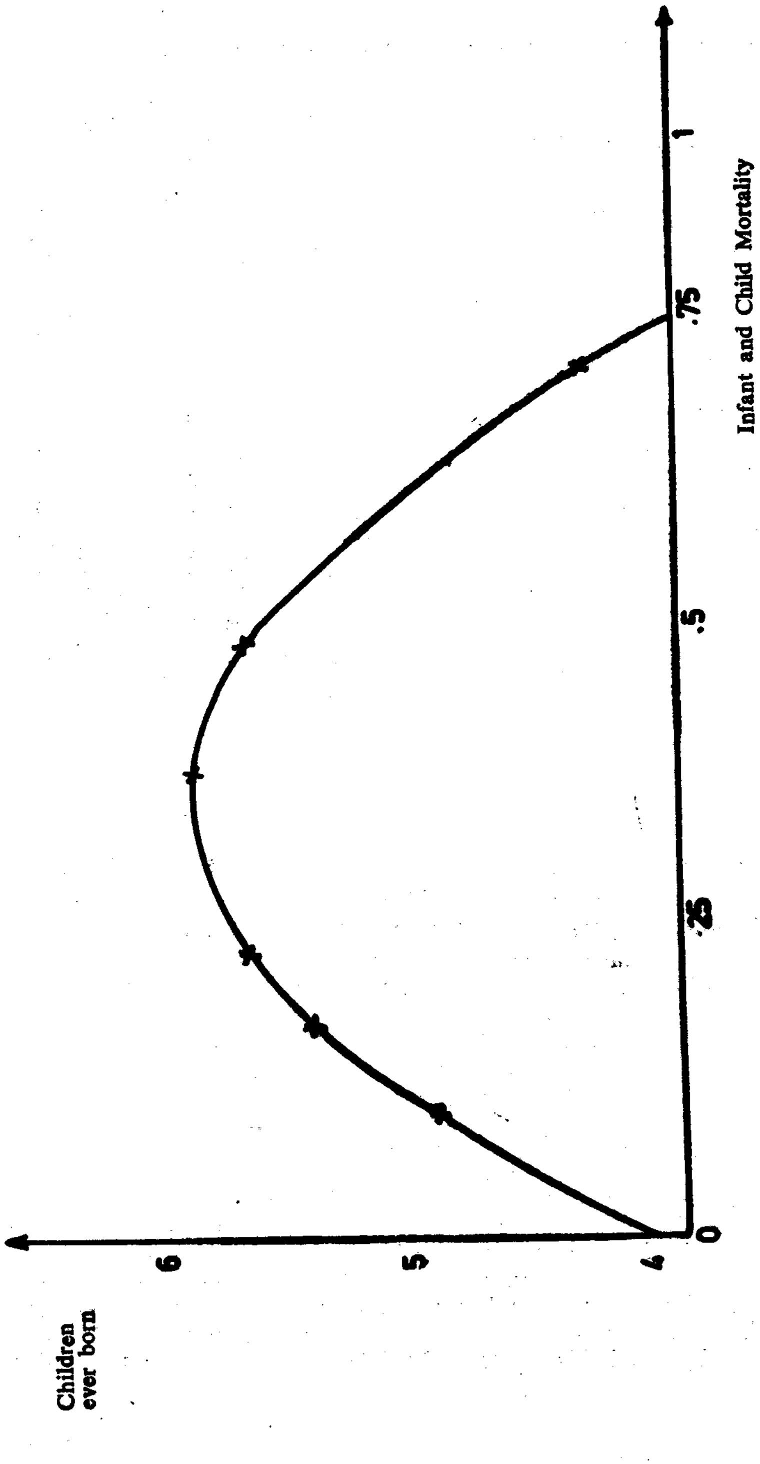

Fig 5
Children ever born and Mortality

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. While there are other methods of exploring the effect of infant and child mortality on fertility (for example, an analysis of birth intervals following an infant or child death), directly assessing the nature of the link between *completed* fertility and mortality would appear to be an impossible task.[8] Further, failure to recognise the nature of the nonbehavioural relationship between mortality and fertility could lead to seriously biased policy conclusions regarding the effect of a decline in mortality on family behaviour.

SUMMARY AND CONCLUSIONS

In this paper, we have tried to demonstrate some of the difficulties researchers face when analysing variations in completed fertility using data relating to women who are still in their childbearing years. Although the purpose of the paper is mainly cautionary, at least one important implication for future research can be drawn from our analysis. With regard to the effect of changes in income on the total number of children a couple has, our results are suggestive of a more complex relationship between fertility and income than has been estimated in many previous regression specifications. Further, although our "best" results (see p. 47) are statistically too weak to serve as a basis for policy conclusions, they do point toward the possibility that a secular rise in income may initially exacerbate the "population problem" in Pakistan by increasing fertility rates, especially among low-income families.

Our analysis of the relationship between infant and child mortality and fertility indicates that many previous studies of this relationship may have produced seriously biased results, and therefore, misleading policy implications. Although it may be possible to develop methodologies to overcome the problem we raise, in our opinion nothing short of complete 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 mortality.

With the increased availability of micro data on individuals 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 recommendation that can be drawn from this study has to do with methodology rather than with 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 a minimum.

---

[8]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 observations will fall in that range.

REFERENCES

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

2. ——————."Age at Marriage, Fertility and Infant-Child Mortality in a Lahore Suburb (Part II)". *The Pakistan Development Review*. Vol. XV, No. 2. Summer, 1976.

3. Becker, Gary S. and Nigel Tomes. "Child Endowments and the Quantity and Quality of Childern". *Journal of Political Economy*. 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 and Fertility: A Policy Review." *Pakistan Development Review*. Vol. XVI, No. 1. Spring 1977.

6. Hamdani, Khalil. "Education and the Income Differential: An Estimation for Rawalpindi City". *The 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.