

# ***Fertility and Marriage in Sri Lanka: Some Insights from Path Analysis***

*Robert W. Hodge  
Naohiro Ogawa*

**NUPRI Research Paper Series No.6**

**October 1981**

**Robert W. Hodge**  
Professor  
Department of Sociology  
University of Southern California

**Naohiro Ogawa**  
Associate Professor  
College of Economics  
Nihon University  
and  
Deputy Director  
Nihon University  
Population Research Institute

## C O N T E N T S

Tables	iv
Figures	iv
Abstract	vi
I. Introduction	1
II. Sri Lanka: A Brief Review	1
III. Data and Measurements	4
IV. Constructing a Causal Model	7
V. Estimating a Causal Model	11
VI. Evaluating the Model	13
VII. Decomposition of Socioeconomic Differentials in Fertility in Sri Lanka	17
VIII. Some Observations on Age at Marriage	21
IX. Conclusion	23
X. Acknowledgements	24
Note	25
Appendices	26
References	28

## T A B L E S

1. Actual and Implied Correlations in a Model of Fertility and Marriage in Sri Lanka .....	15
2. Decomposition of the Relationship Between Type of Place of Childhood Residence (U) and Fertility (F) .....	18
3. Decomposition of the Relationship Between Education (E) and Fertility (F) .....	19

## F I G U R E S

1. A Map of Sri Lanka Indicating the Six Zones .....	2
2. A Model of Fertility and Marriage in Sri Lanka .....	10

## A B S T R A C T

---

By drawing upon data collected in the Sri Lanka Fertility Survey conducted in 1975, this study attempts to clarify, through path analysis, how age at first marriage and other factors affect fertility. The causal model developed in the present study includes both nominal and interval variables as well as an unmeasured variable. Although we have imposed in the model the extreme assumption that no net socio-economic differentials in fertility were present in Sri Lanka, the model nonetheless performs quite satisfactorily. Despite the fact that it is a simplification of reality to assume that the entire causal import of socio-economic factors upon fertility is channelled through age at first marriage and region, these are undoubtedly major vehicles through which socio-economic forces affect cumulative fertility in Sri Lanka.

---

## I. Introduction

Over the past two decades a number of Asian nations have experienced fertility declines of various degrees. These declines have varied in magnitude: some like those in Hong Kong (Freedman et al., 1970; Chan, 1976) and Japan (Ohbuchi, 1976) have been substantial and sustained, while others, such as those apparently underway in Thailand (Knodel and Debavalya, 1978) and Sri Lanka (Fernando, 1976), have been more modest in scope with respect to both their magnitudes and time horizon. The empirical documentation of these fertility declines is beyond reasonable dispute, but the causal mechanisms responsible for them are less than certain. In the case of Hong Kong, for example, both family planning and economic development were closely associated with the decline in fertility and various authors have argued that causal primacy should be assigned to one rather than the other of these factors (Wat and Hodge, 1972; Freedman, 1973).

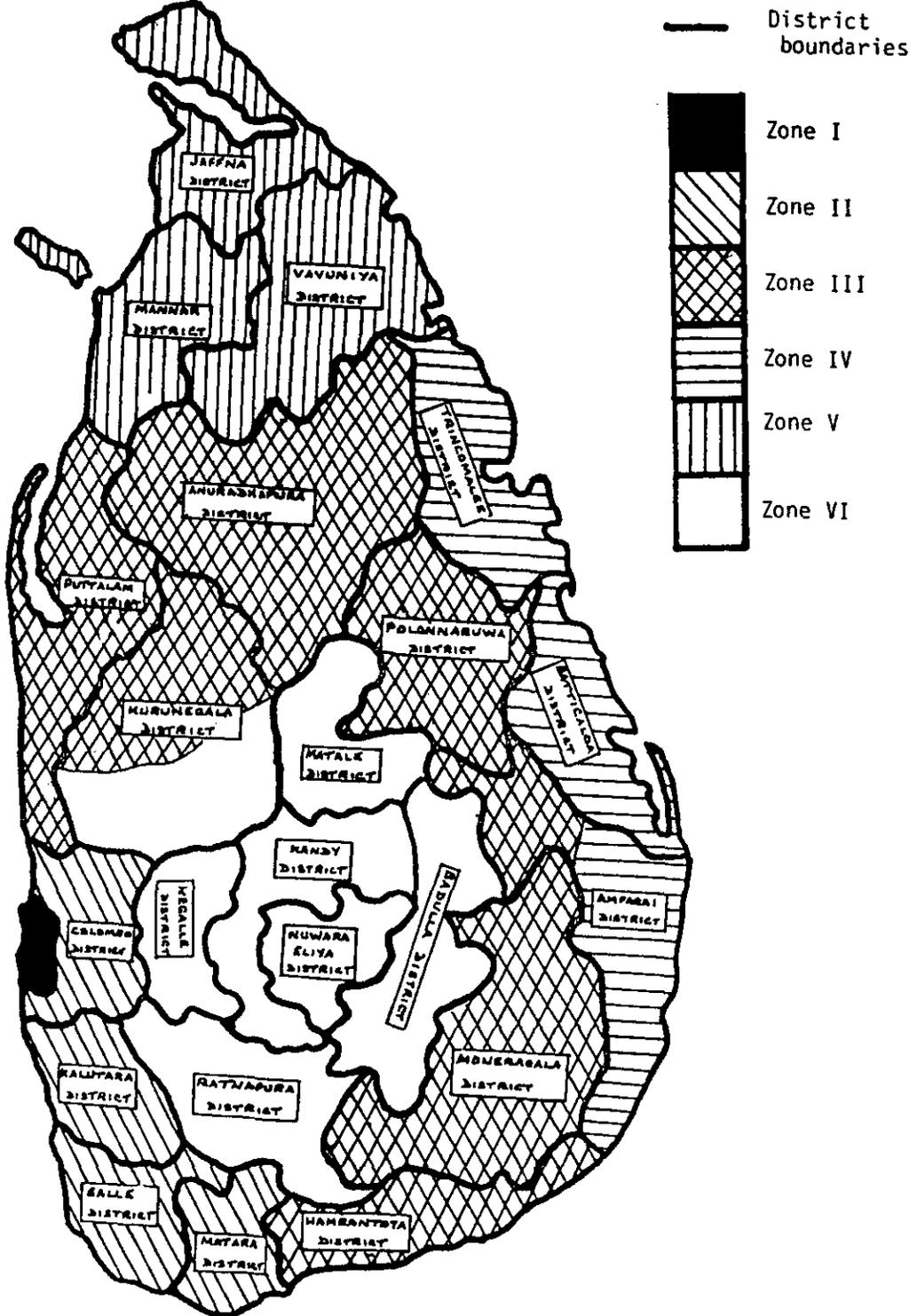
Family planning activities and economic expansion by no means exhaust the factors implicated in fertility swings, as recourse to the demographic history of the West makes clear. There is emerging evidence that marital patterns and, particularly, shifts in the age at first marriage may be responsible for a significant fraction of the fertility declines observed in some Asian nations (Cho and Retherford, 1973). Sri Lanka is one of those nations.

In order to understand how age at first marriage and other factors affect fertility it is necessary to have a causal model of the way they are implicated in the process of fertility determination. Our purpose in this paper is to conjecture and evaluate such a model, drawing upon data collected in the World Fertility Survey for Sri Lanka. We begin with a brief review of the setting for our analysis.

## II. Sri Lanka: A Brief Review

As shown in Figure 1, Sri Lanka is an island country, located to the Southeast of India with a territory size of slightly more than 25,000 square miles. According to the 1971 Sri Lankan Census, its population size is approximately 13 million, with four major ethnic groups: the Sinhalese, 71 percent of the total population, Sri Lanka Tamils, 11 percent, Indian Tamils, 9 percent, and Sri Lankan Moors,

Figure 1. A Map of Sri Lanka Indicating  
The Six Zones\*



\* This map has been copied from Figure 2.2 in the  
First Report of the Sri Lanka Fertility Survey.

7 percent.

This island country is comprised of two geographically distinctive areas: the Dry Region and the Wet Region. The Dry Region, which has a rainfall of less than 75 inches a year, spans the southernmost part of the island to the northwestern part, circling the South Central Hill Country on the east. The rest of the island is regarded as "the Wet Region". Compared to the rest of the island, Colombo, the highly urbanized capital of the country, exhibits markedly different socio-economic characteristics.

Sri Lanka, being pre-dominantly an agricultural society, exports rubber, tea, and coconut products to foreign markets. In 1975, its per capita gross national product amounted to \$150. As opposed to its early stage of economic development, Sri Lanka has attained an advanced level of social development. Due to its strong national emphasis on social development, the educational level of the Sri Lankan population is very high in comparison to other developing countries in Asia. In 1971, 85 percent of males and 70 percent of females above age 10 were literate.

Demographically, the crude death rate declined dramatically from 19.9 to 11.2 between 1941-1945 and 1951-1955, as a result of the island-wide malaria eradication drive. Since then, the crude death rate has been dropping gradually, thus reaching an average of 8 per thousand over the period of 1971-1974. As regards fertility, the crude birth rate fell from 38.1 in 1951-1955 to 36.5 in 1956-1965, and continued to decrease at a steady pace to 28.7 in 1971-1974. As demonstrated elsewhere (Ogawa and Rele, 1981), the major source of this fertility reduction was the change in the nuptiality pattern. More specifically, from 1963 to 1971, the proportion of women married fell from 14.7 to 10.4 percent for the age group 15-19, from 57.6 to 45.9 for the age group 20-24, and 81.0 to 73.4 for the age group 25-29. Although the inception of the National Family Planning Program in 1965 gave rise to a considerable decrease of marital fertility in the late 1960s, its contribution to the decline in the birth rate was not as pronounced as that of changes in the marital structure.

In examining Sri Lanka's recent changes in the marital pattern, one should bear in mind a variety of cultural, socio-economic and ecological elements. One of the studies (Ryan, 1952) shows that the matchmaking process involves five fundamental rigid criteria and

three balancing factors; these five criteria include ethnicity, caste, age differentials, bride's virginity and horoscope-compatibility, while the balancing factors are girls' families dowry power, security and occupation of the groom, and the status of family lines. Moreover, the spouse selection process differs with ethnicity. For instance, in Hinduism a girl should marry before she reaches puberty. In the Muslim tradition, the status of women is considerably lower than that in other religious traditions. In addition to these ethno-religious differentials, the match-making process becomes further complicated by differences in the dowry system. For both Sinhalese and Sri Lankan Tamil, the dowry reflects the financial status of brides and their families. For this reason, young women tend to work to alleviate the economic strains placed upon their parents in the accumulation of financial resources for the dowry. Among Muslims, instead of a dowry, the "Mehr" practice, which is paid by the groom to his bride, is prevalent. Because of this tradition, Muslim males show relatively high age at marriage and Muslim females, low age at marriage.

As distinct from other South Asian societies, Sri Lanka has been traditionally open to the education of women, resulting in the massive expansion of education in the past three decades, which in turn, has facilitated the rising status of women. Furthermore, women's age at marriage and female participation in the labor force have a direct relationship. This relationship has been pronounced in Sri Lanka. Nonetheless, women on the estates are rather an exception. They are inclined to marry young regardless of their pre-marital work status.

Because of the recency of its National Family Planning Program, age at first marriage may be the best single predictor of fertility in Sri Lanka (Ogawa and Rele, 1981). It should be noted, however, that subfecundity is the other factor worth considering in the Sri Lankan context. The low fertility of women on the estates may be partially explained by this subfecundity hypothesis. (In Figure 1, most of the estates are contained in Zone VI.)

### III. Data and Measurements

The Sri Lankan segment of the World Fertility Survey was com-

pleted in 1975. Interviews were completed with 6,870 ever married women under age 50, using the basic instruments developed for the World Fertility Survey program. The present analysis is restricted to a truncated set of 4,774 ever married women. This truncation is necessary because those young women who were still unmarried at the time of survey were not included. As a solution to the bias of the data set, we have truncated the data so as to make the exposure of risk to marriage for all women comparable. Following the earlier analysis (World Fertility Survey, 1978), we have used the pivotal age of 25; those who were at age 25 and above, and who were married before the age 25 have been selected for analytical purposes of the present study.

Although other efforts at causal modelling using World Fertility Survey data have been completed, one is well advised to remember that causal modelling of the process of family formation was not the primary purpose of these inquiries. The data are at once more limited --with respect to the determinants of fertility--and more extensive with respect to the demographic histories of individual women than would be data sets whose primary purposes were to implement the estimation of causal models of fertility and related processes. As in any secondary analysis, one is limited in using the World Fertility Surveys for purposes of causal modelling by virtue of the data they contain.

Given these limitations in the data, it is self-evident that one could not possibly hope to construct a complete model of the process of fertility and family formation. The best one can hope for is to hit on the correct causal model relating the variables available for analysis. Put otherwise, it is doubtful that any equations one is likely to estimate will in fact correspond to the true structural equations; they will almost surely be, at best, quasi-reduced forms. Nonetheless, important insights can be obtained about how a basic, if incomplete, set of variables are interrelated by examining them in a causal context. The history of the use of structural equation methods to examine processes of status attainment (Blau and Duncan, 1967; Sewell et al, 1970) and to model economic changes (Tinbergen, 1939; Klein and Goldberger, 1955; Ball, 1973) would provide substantial documentation for this claim.

Despite these reservations, the Sri Lankan segment of the World

Fertility Survey contains measures upon many of the social, economic, and demographic factors thought to govern differential fertility in that country. From the available indicators, we have selected ten for analysis herein. They and the symbols used to represent them are as follows: (1) U, type of childhood residence; (2) Y, year of birth; (3) R, race; (4) C, religion; (5) E, years of school completed; (6) W, work status prior to first marriage; (7) H, age at first marriage; (8) D, months since first marriage; (9) A<sub>1</sub>, region of current residence; and (10) F, number of children ever born.

Of the foregoing variables, half--U, R, C, W, and A<sub>1</sub>--are nominal, rather than interval or ratio variables. Inclusion of such categorical indicators in a path analysis is facilitated if they can be indexed along a continuous scale. (This might seem like a requirement, but strictly speaking it is not. All but one of the categories of the qualitative variables could be treated as a separate dummy variable, but the resulting analysis would be clumsy to say the least.)

There are a variety of ways in which nominal variables can be indexed so they can be treated as if they were measured at an interval or ratio level. None of these methods is wholly satisfactory, for in the final analysis they are what they are--nominal variables. Nonetheless, if one wants to get a picture of the overall causal structure, recourse to one or another method of indexing greatly simplifies the subsequent analysis. For present purposes, we chose to index the categories of the nominal level variables according to the mean number of children ever born to the women in them. This method of scoring was not capricious. It maximizes the zero order association of each of the nominal variables with the penultimate dependent variable in our model, viz., children ever born. This gives each one of the nominal level indicators its best possible starting place.<sup>1/</sup>

Given the method by which the nominal variables were scored, it is important to remember that the analysis which follows is fertility biased, that is, the relationships observed between pairs of nominal variables and the remaining variables in the analysis reflect that part of the nominal variables which is related to fertility. Consequently, the associations observed between each pair of nominal variables and between each nominal variable and each of the remaining

variables is necessarily less than the maximum associations which would be found by relating the nominal variables to each other via canonical correlation of dummy variables representing the categories of each nominal variable or by regressing each of the interval or ratio variables on these same dummy variable representations of the categories of the nominal level variables. The categories of each of the nominal level variables, the numbers of women in them, and their mean number of children ever born are given in Appendix Table 1. The means and standard deviations of all the variables are given in Appendix Table 2.

#### IV. Constructing a Causal Model

The first step in constructing a causal model is identifying the complete or partial causal ordering of the variables involved. This ordering may be based on the way the variables are temporarily sequenced, theoretical considerations, or findings from previous research. In the present case, the causal ordering of the variables appears to be reasonably unambiguous. Four of the variables--type of childhood residence, year of birth, race, and religion--are for all practical purposes ascribed characteristics over which the respondent has little or no control. This is especially the case in view of the fact that very few respondents are likely to have changed their religious affiliation over their lifetimes. Consequently, we simply treat U, Y, R, and C as exogenous variables--given in our model just as they are given to our respondents. We regard schooling (E) and work status before marriage (W) as the first two endogenous variables in our model. They are doubtlessly related--simultaneously; for simplification we here simply presume they are causally un-ordered and unconnected directly. The next variable to appear in the causal order is age at first marriage (H), which is followed by duration of marriage (D) and region of current residence ( $A_1$ ). We do not postulate a causal ordering of D and  $A_1$ , which seems to us ambiguous. Regional geographic mobility within marriage may not be large, as we suggest below. Since the survival of marriages is in part dependent upon the socio-economic environment in which they are conducted, one could argue that current region, which reflects one's socioeconomic environment to some extent, should be regarded as causally prior to marital duration. However, divorce may occasion

interregional geographic mobility; to the extent this occurs, current region is properly regarded as a consequent rather than a cause of marital duration. We beg the question by imposing no causal order upon this pair of variables. The final and penultimate variable in our causal sequence is, of course, number of children ever born (F). The causal order we have postulated appears to us reasonable, though not totally immune from dispute, since it is based in large measure upon the overwhelming numerical preponderance of the ways individual life cycles are sequenced. *But there are some women who do complete their education in marriage and it would not be preposterous to regard work status before marriage and age at first marriage as simultaneously determined.*

Having settled upon a causal order, one may proceed to estimate recursive models such as the one we have postulated in a straightforward and mechanical way. The procedure is often referred to in the literature as theory trimming and typically involves a two-step estimation procedure. First, each endogenous variable is regressed on every exogenous and endogenous variable which is causally prior to it. The coefficients obtained in the resulting regressions are inspected for their statistical and/or substantive significance. A second round of estimation is then undertaken, with the variables deemed insignificant deleted from the equation. Voila! Estimation is complete and one proceeds to interpretation.

We believe, however, that such mechanical determination of what effects are left in or out of models invites little, if any, room for the operation of speculation and theoretical considerations, save in the determination of the causal order. As much, if not more insight and understanding can often be gleaned from the estimation and evaluation of conjectural models which reflect a particular theoretical position or a substantive state of affairs of particular interest. One seeks not so much to estimate or find the correct model, as to show what the model would be if a particular substantive condition or a particular theory were true. Evaluating the model becomes of paramount importance, because that evaluation reveals the extent to which one's theory or a postulated substantive arrangement can be maintained in the light of one's data. We have chosen to opt for this approach, rather than mechanically producing a model by pruning and, if necessary, re-pruning the equations until all the

path coefficients are significant.

The model we have postulated is shown in Figure 2 in the form of a path diagram. (Estimates of the path coefficients are also shown in the figure, but these should be ignored for the moment.) Several features of the model deserve comment, but the most striking one is obvious: there are no net socioeconomic differentials in fertility in Sri Lanka! Instead, the conjectured model forces the entire gross impact of childhood residence, race, religion, education, and premarital work status to be channelled through age at first marriage, region of residence, or both. Of these two factors, age at first marriage is clearly the most pivotal, owing to the substantial indirect way it proves to be linked to fertility.

The state of affairs described above has been built into the conjectured model chiefly by virtue of the simple equation which makes fertility a function only of exposure and environment (as reflected in current region of residence). Exposure, or duration of marriage, is then treated as a function of year of birth and age at first marriage alone. This latter equation is virtually an identity, since for currently married, once married women,  $D$  exactly equals  $1975 - Y - H$ . Only instances of widowhood and divorce make this equation stochastic rather than deterministic.

There is one other feature of our conjectured model which is worthy of note. We have introduced an unmeasured variable, region of residence at marital age ( $A_0$ ), which is allowed to affect age at first marriage. The reason for introducing this hypothetical variable is rooted in the rather substantial observed correlation of  $-.2360$  between region of current residence and age at first marriage. It makes little substantive sense for current residence to be affected by age at first marriage, the causal direction of influence implied by the temporal ordering of these two variables. However, it does make a great deal of sense that region should affect age at first marriage, if for no other reason than that the marriage market conditions of regions are likely to differ. But the relevant regional variable affecting first marriage is not where one presently lives, but where one lived at the time of first marriage. Consequently, we have introduced the hypothetical variable, region of residence at marital age to resolve the dilemma posed by a correlation between current region and age at first marriage that is too large to ignore.

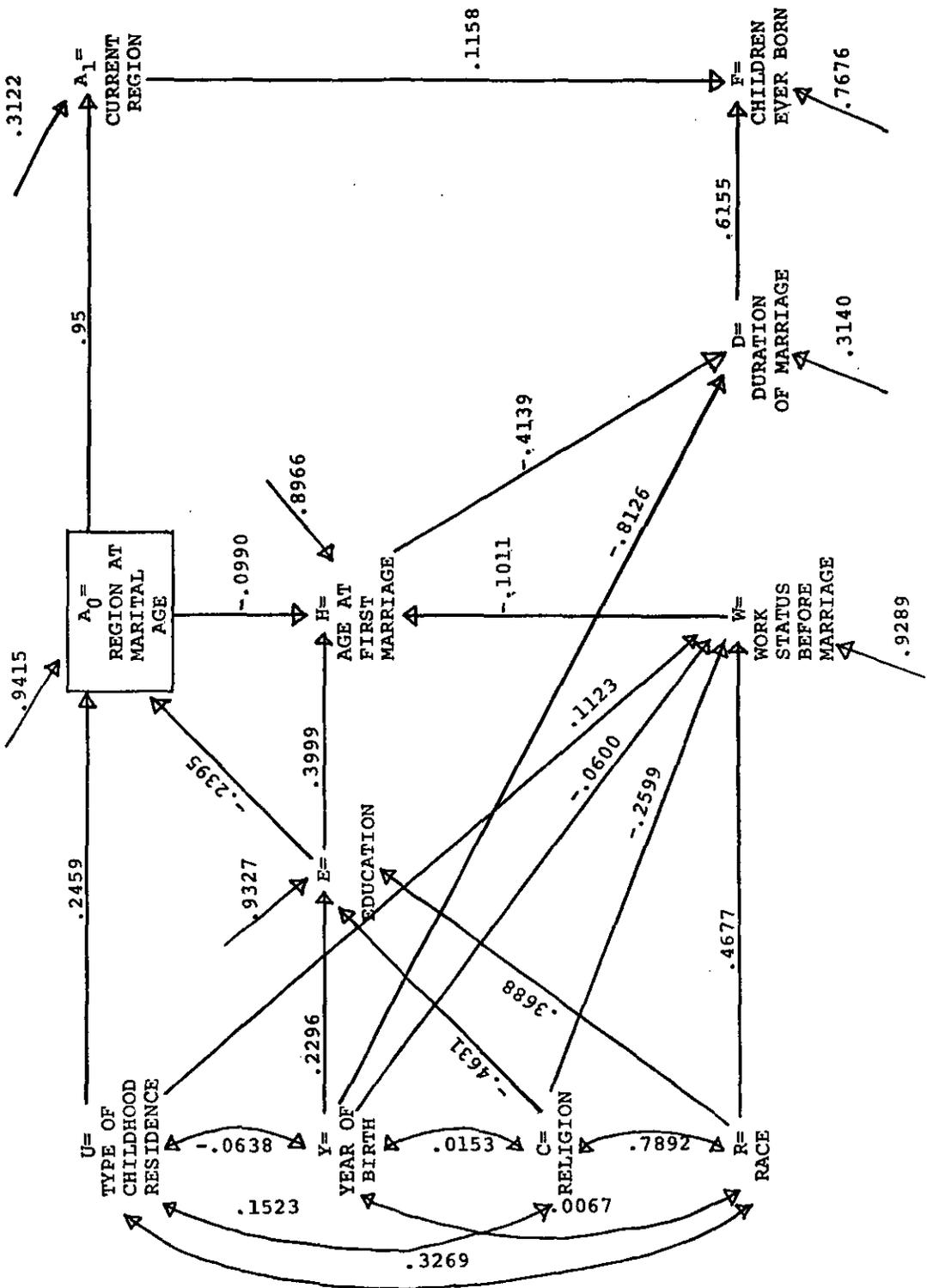


Figure 2. A Model of Fertility and Marriage in Sri Lanka

The remaining portions of the conjectured model demand no explanation at this juncture. They merely reflect how the exogenous, ascribed characteristics operate partially to determine educational attainment, as well as work status and regional residence at marital age. In addition, education is allowed to enter the equation for region at marital age. Save for the effect of year of birth on marital duration, the ascribed, exogenous variables are predicated to have no direct impact beyond their influence on education, region at marital age, and work status before marriage--the three factors entering the equation for age at first marriage. This last part of the model represents no particularly strong substantive assumptions, for it simply reveals the operative mechanisms by which these background factors indirectly affect fertility via region and age at first marriage.

#### V. Estimating A Causal Model

The model postulated in Figure 2 is overidentified and for that reason there are numerous possible ways it might be estimated. However, where feasible we have utilized ordinary least squares regressions of the endogenous variables upon the factors which directly influence them, since this method produces the lowest standard errors of the estimated coefficients (Goldberger, 1970). The equations for F, D, W, and E were estimated in this way.

Estimation of the remaining equations cannot be accomplished by ordinary least squares because they involve the unmeasured variable  $A_0$ . Indeed, they cannot be estimated at all without making a critical assumption about the relationship between region of current residence and region at marital age. The Survey of Fertility in Sri Lanka does not contain any data on interregional migration and we could not obtain data of this type from alternative sources. We do, however, have tables showing interregional lifetime and intercensal migration for the Philippines and another table pertaining to interregional migration in Malaysia. For other purposes we had undertaken canonical analyses of these data sets. The first canonical correlation associated with each of them was without exception well in excess of .9; discussion with a few colleagues in South Asia gave us no reason to reject this assumption. Consequently, we simply imposed

the value of .95 on the association between region of current residence and region at marital age. The actual value is surely a bit higher or a bit lower than this, but the figure we have utilized is almost certainly in the same ballpark.

Estimating the equations for  $A_0$  and for H is a straightforward exercise in indirect least squares once the equation for  $A_1$  is fixed by the assumption in the foregoing paragraph. We have

$$A_1 = .95(A_0) + e \quad (1)$$

$$\text{and } A_0 = \lambda(U) + \chi(E) + f. \quad (2)$$

Substituting Eq. 2 into Eq. 1 gives a partially reduced form for  $A_1$ :

$$A_1 = .95\lambda(U) + .95\chi(E) + (.95f + e). \quad (3)$$

Regressing  $A_1$  on U and E, we find  $.95\lambda = .23357$  and  $.95\chi = -.22751$ , which yield the reported values of  $\lambda$  and  $\chi$ .

Estimating the equation for H is equally straightforward. We have

$$H = \alpha(A_0) + \beta(E) + \gamma(W) + h. \quad (4)$$

Substituting Eq. 2 into Eq. 4 yields a partially reduced form for H:

$$H = \alpha\lambda(U) + (\beta + \alpha\chi)(E) + \gamma(W) + (\alpha f + h). \quad (5)$$

Regressing H on U, E, and W we find  $\alpha\lambda = -.024343$ ,  $\beta + \alpha\chi = .42364$  and  $\gamma = -.10108$ . These are readily solved to find the reported values of  $\alpha$  and  $\beta$ , since the values of  $\lambda$  and  $\chi$  have already been estimated.

The description of the estimation of the model at hand is now virtually complete, since only the derivation of residual paths remains. In a fully recursive model in which all possible paths are present, they may be obtained by simply taking the square root of  $1 - R^2$ . This, however, is inappropriate in overidentified models like the present one, since it can yield results that imply the correla-

tions of some of the endogenous variables with themselves are greater or less than one. In the present case, we have estimated the residual paths by working out the correlation of each endogenous variable with itself. This yields an estimate for each of the residual paths when the parameter estimates already at hand are substituted into the resulting equations. In the case of the residual paths for E, W, and A<sub>1</sub> the results are identical to those obtained in the usual way, i.e., by taking the square root of  $1 - R^2$ . Elsewhere, however, the results are not equivalent. Our estimates, however, are consistent, since they yield estimates of the residual paths which force, within rounding error, each endogenous variable to have a correlation of unity with itself.

#### VI. Evaluating The Model

The estimates of the model reported in Figure 2 contain, with a single exception, no surprises. The massive impact of marital duration observed throughout the developing world is here observed in Sri Lanka. Education is obviously a signal driving force in the determination of age at first marriage and there is a substantial increment in years of schooling completed with each successive cohort, a point which suggests that the present downward trend in Sri Lankan fertility will almost surely continue if educational opportunities for women expand. The remaining coefficients, with the exceptions noted below, are of the expected sign. All are statistically significant by the conventional criteria, even though some are numerically small.

The effects of race and religion on education and work status before marriage are, however, confusing and inconsistent. Women in religious groups with high fertility levels complete, on the average, fewer years of schooling than those in religious groups with lower fertility levels. That is as expected, but these same women in religious groups with high fertility levels are also likely to wind up in work statuses before marriage that are associated with lower fertility. In and of themselves, these contrasting effects are not inexplicable, for they are consistent with a widespread dowry system in non-Muslim groups which often requires that young women work in order for their families to raise the funds necessary to secure their marriages. That explanation, however, is hard to believe, because the pattern observed for women in high fertility racial groups is

just the reverse: as expected they are more likely to wind up in work statuses before marriage associated with high fertility, but they are also likely to complete more rather than fewer years of schooling. In sum, the effects of religion and race--as observed in the model--virtually cancel each other out, leaving little joint effect of the two on either education or work status. This is in agreement with some of the earlier research findings (Ogawa and Rele, 1981). The substantial multicollinearity between race and religion is the likely villain responsible for the inconsistent results observed when one attempts to split out the separate impacts of these two variables.

The fundamental theorem of path analysis states that:

$$r_{ij} = \sum_k p_{ik} r_{kj}$$

where either the *i*th and *j*th variables are causally unordered or where the *j*th variable is causally prior to the *i*th (Wright, 1921). The estimating equations used in path analysis are often derived by explicit use of this formula. However, the formula can also be used to derive the correlations which are implied by an estimated path diagram. To find the implied association between the *i*th and *j*th variables, the estimated values of the  $p_{ik}$ 's are substituted on the right. Also, one uses the implied, rather than the actual values of the  $r_{kj}$ 's on the right. Sometimes, of course, the implied correlations will necessarily equal the actual ones owing to the estimating strategy and the structure of causal effects postulated in the model. However, often this is not the case, even though OLS has been used for estimating all portions of a model. This occurs when the model is overidentified, i.e., some possible effects have been set equal to zero.

We have worked out the correlations between the variables implied by the model shown in Figure 2. These, together with the actual correlations observed between the variables, are shown in Table 1. A comparison of the implied and actual correlations provides an informal basis for assessing the performance of the model and for identifying features of the original data which are inadequately captured by the model. Wide discrepancies between the actual and implied associations indicate the model should be revised or abandoned. Consistency

Table 1. Actual and Implied Correlations in a Model of Fertility and Marriage in Sri Lanka

Variables	Variable										
	U	Y	R	C	E	W	H	D	A <sub>1</sub>	F	
	Actual Correlations (Above Diagonal); Implied Correlations (Below Diagonal) *										
U	1.0	-.0638	.3269	.1523	.0194	.2294	-.0393	.0779	.2292	.1183	
Y	...	1.0	.0067	.0153	.2250	-.0680	.0785	-.8451	-.0498	-.4782	
R	...	...	1.0	.7892	.0049	.2989	-.0726	.0177	.1292	.0875	
C	...	...	...	1.0	-.1685	.1253	-.0955	.0164	.1200	.0799	
E	.0354	...	...	...	1.0	.1078	.4123	-.3489	-.2230	-.3041	
W	...	...	...	...	.0366	1.0	-.0610	.0867	.0590	.0957	
H	-.0325	.1037	-.0360	-.0879	.4190	-.0912	1.0	-.4776	-.2360	-.3851	
D	.0633	-.8555	.0095	.0239	-.3563	.0930	-.4982	1.0	.0675	.6233	
A <sub>1</sub>	.2255	-.0661	.0739	.0752	-.2193	.0453	-.1868	.1308	1.0	.1573	
F	.0651	-.5342	.0144	.0234	-.2447	.0625	-.3283	.6306	.1963	1.0	
A <sub>0</sub>	.2374	-.0696	.0778	.0792	-.2308	.0477	-.1961	.1377	...	...	

\* (Implied correlations are given only when the method of estimation allows their value to differ from the actual.)

of the implied and actual correlations reveals only that the model describes the data well; it does not mean one has found the correct causal model.

A comparison of the actual and implied correlations reported in Table 1 reveals, despite the rather extreme substantive assumptions which were built into the model, a surprisingly good fit between the two sets of figures. Nonetheless, examination of the actual and implied correlations, pair by pair, turns up some discrepancies indicative of model failures. The model misses the positive association between education and work status before marriage by more than .07. The positive association between education and premarital work status means that more educated women are less likely to work before marriage. Realistically, premarital work status and education should probably be regarded as simultaneously determined, but there is little hope, in our judgement, of providing a satisfactory estimate of that simultaneity with the data in hand. We eschewed an arbitrary and conceptually erroneous, recursive solution by forcing education into the equation for work status or vice versa. This leaves us with a shortfall in the implied association between work status and education.

The model also underestimates the association of both race and religion with current region of residence by an amount of about .05. One could easily remedy this failure to capture the regional concentration of racial and religious groups by allowing one or both of these variables to influence region at marital age. The model also understates the association of age at first marriage with region, while overstating the association between region of current residence and duration of marriage.

The model misses most of the association between both race and religion with fertility and understates (in absolute terms) by about .06 the full association of both age at first marriage and education with children ever born. At the same time, the model overstates the associations between fertility and both duration of marriage and region of current residence. All of these problems would be remedied in some degree by allowing age at first marriage to affect fertility directly.

Despite the discrepancies noted above, the model still performs remarkably well in view of the extreme substantive assumption built

into it. The results to this point leave little doubt that a substantial fraction of the socioeconomic differentials in fertility in Sri Lanka are, indeed, channelled through age at first marriage and/or regional location. We assess the full extent to which this is the case in the following section.

#### VII. Decomposition of Socioeconomic Differentials in Fertility in Sri Lanka

The rules of path analysis may be used to decompose the implied correlations into the constituent causal and spurious linkages which make them up. In Table 2, we have made such a decomposition of the gross association between type of childhood residence and fertility. In this and subsequent decompositions we have lumped the spurious portions of the implied associations together in a single figure, since the constituent elements of the spurious component are of little interest in themselves.

Only a little over half of the gross association between type of childhood residence and fertility is captured by the association implied by the model. This residue represents model error and has three potential components: (1) a possible direct effect of childhood residence on fertility, (2) additional indirect effects which are due to missing linkages elsewhere in the model, and (3) additional spurious elements due to other missing linkages. There is no way to know how this residue is itself divided into these components. At worst, the entire residue represents a direct impact of childhood residence on fertility, but that is very unlikely owing to the location of this variable in the model.

Of the portion of the association between childhood residence and fertility which can be captured by the model, about half is spurious and half indirectly causal. The causal portion is overwhelmingly transmitted via the regional variables, though 15 percent does flow through age at first marriage via region or work status before marriage. In sum, then, at least 30 percent of the gross association between childhood residence and fertility is channelled through region and age at marriage. This represents two-fifths of the part which is not demonstrably spurious and almost certainly more of the part which is actually non-spurious.

Table 2. Decomposition of the Relationship Between  
Type of Place of Childhood Residence (U)  
and Fertility (F).

Component	Magnitude	Percent of Gross Association	Percent of Implied Associ- ation
Gross Association	<u>.1183</u>	<u>100.0</u>	...
Total, Implied Association	.0651	55.0	100.0
A. Indirect Causal Impact	<u>.0362</u>	<u>30.6</u>	<u>55.6</u>
1. Via Regional Variables	.0271	22.9	41.6
2. Via Region and Age at First Marriage	.0062	5.2	9.5
3. Via Work Status and Age at First Marriage	.0029	2.5	4.5
B. Spurious (Residue of Implied)	.0289	24.4	44.4
Difference (Gross less Implied)	.0532	45.0	...

Table 3. Decomposition of the Relationship Between Education (E) and Fertility (F).

Component	Magnitude	Percent of Gross Association	Percent of Implied Association
Gross Association	<u>-.3041</u>	<u>100.0</u>	...
Total, Implied Association	-.2447	80.5	100.0
A. Indirect Causal Impact	<u>-.1342</u>	<u>44.1</u>	<u>54.8</u>
1. Via Age at First Marriage	-.1019	33.5	41.6
2. Via Region at Marital Age and Age at First Marriage	-.0060	2.0	2.5
3. Via Regional Variables	-.0263	8.6	10.7
B. Spurious (Residue of Implied)	-.1105	36.4	45.2
Difference (Gross less Implied)	-.0594	19.5	...

The decomposition of the correlation between education and fertility is given in Table 3. Education is clearly more closely associated with fertility than is childhood residence and the model performs much better with respect to this nexus. Four-fifths of the gross association between education and fertility are captured by the implied correlation. As was the case with childhood residence, a little more than half of the implied association is indirectly causal and a little less than half is spurious. Of the causal part, about four fifths-- $(44.1)/(54.8) = .805$ --passes through age at first marriage. Even if the entire difference between the actual and implied associations represented a direct causal impact of education on fertility--a highly unlikely circumstance, the indirect causal influence through age at first marriage and region would be nearly twice as large. In sum, nearly half of the gross association between fertility and education is passed through either region or age at first marriage. This amounts to almost 70 percent of the fraction of the gross association which is not demonstrably spurious.

Turning our attention to the relationship between work status and fertility, we again find a shortfall in the implied association relative to the total association. The former captures only about two-thirds of the latter. Much of the implied association is spurious. The one indirect causal linkage admitted by the model through first marriage amounts to  $(.6155)(-.4139)(-.1011) = .0258$ , or about two-fifths of the implied and one-quarter of the total. Even if the entire difference between the implied and actual association represented direct and/or indirect causal forces, premarital work status would still exert very little impact upon fertility.

Assessing the relationship between fertility and either race or religion is made difficult by virtue of their multicollinearity and their inconsistent effects. We may, however, note that the model fails to account for their observed association with fertility. The coefficient of determination linking fertility to race and religion alone is .0079; if we compute the same coefficient of determination using the correlations implied by the model, we obtain .00059, a value less than one-tenth as large. But the numbers are miniscule, no matter how one looks at them. Race and religion just do not have a great deal to do with fertility in Sri Lanka. Again, this result agrees with some of the earlier research findings (Ogawa and Rele,

1981).

This exercise in statistical decomposition of gross associations reveals several points of interest. First, it highlights some of the failures in the model: there are residues in the total associations linking socioeconomic variables to fertility in Sri Lanka which the model does not explain. Doubtless, contrary to the model specification, there are some net impacts of socioeconomic variables upon fertility. While this may be true, we may also observe, second, that the magnitudes involved are at best quite small relative to the estimated effects in the model. That is basically why the model performs so well.

#### VIII. Some Observations on Age at Marriage

The model at hand admits of only one causal impact of age at marriage on fertility: the indirect one via duration of marriage. The force of this causal connection is  $(.6155)(-.4139) = -.2548$  and amounts alone to  $(100)(-.2548)/(-.3851) = 66.2$  percent of the gross association between fertility and age at marriage. However, the model only captures  $(100)(-.3283)/(.3851) = 85.3$  percent of the total association between fertility and age at marriage. While this figure is impressive, it is not so large as to rule out the possibility that age at first marriage exerts a direct causal impact upon fertility.

We explored the possibility suggested above by including age at first marriage in the fertility equation. Doing this, we found that  $P_{FD}$  and  $P_{DA1}$  fell slightly to .5749 and .0979, respectively. However,  $P_{FH}$  was -.0874, which implies that the total direct and indirect causal impact of age at first marriage is  $(-.0874) + (-.4139)(.5749) = -.3253$ . Thus, by the simple stratagem of putting age at first marriage in the equation for fertility, the causal import of marital age is raised by a factor of  $(-.3253)/(-.2548) = 1.28$ . Thus, without changing the basic spirit of the present model, a single simple change will raise the fraction of the gross associations of socioeconomic variables which are channelled through age at first marriage by 25 percent! The cost of achieving this gain would be to lower the amounts passed through region of current residence by a factor of  $(.0979)/(.1158) = .85$ , or about 15 percent. This is, indeed, a small price to pay since most of the action is already through age at first

marriage rather than region.

One further comment seems warranted. Although the model in hand contains some of the major variables thought to influence age at first marriage, they fall far short of determining the timing of marriage in the chronological life cycle. Any doubt about this is clearly quelled by simply inspecting the order of magnitude of the residual path leading to age at first marriage in Figure 2. This means there is considerable leeway for affecting age at first marriage other than by manipulating the variables known to influence it.

Considerable effort has gone into devising ways of lowering fertility by (a) expanding educational and job opportunities for women, (b) stimulating economic development, and (c) reducing marital fertility via family planning programs, among others. We know, however, of no serious effort to invent policies geared at raising the age of first marriage. There are probably several reasons for this, doubtlessly one of them being the tendency to think of age at first marriage as a culturally prescribed norm. But in Sri Lanka, age at first marriage is far from a norm, in the frequency sense. The mean age at marriage is 17.7, but its standard deviation is a very substantial 3.5. The mean number of children ever born among ever married women in Sri Lanka is 4.9, with a standard deviation of 2.8. Results presented elsewhere in this section indicate that a slight modification of this model implies that raising the age at first marriage in Sri Lanka by a standard deviation would result in lowering fertility by about one-third of a standard deviation, which is nearly equivalent to one whole child since  $(.3253)(2.8) = .91$ . Thus, at their respective means, a modest modification of the model in hand implies that a  $(100)(3.5)/(17.7) = 19.7$  percent rise in the age at first marriage would yield roughly a  $(100)(.9)/(4.9) = 18.4$  percent reduction in fertility. Figures like this, of course, are at best crude and imprecise. Nonetheless, they suggest that there would be considerable payoff to any policy successful at enticing young women to withdraw themselves from the marriage market. This is not the place to speculate about what such policies might be but just launching some surveys designed to reveal the conditions under which young men and women enter the marriage market and contract for marriages within it would be a useful beginning for either formulating such policies or deciding that there are no financially feasi-

ble, viable ones.

## IX. Conclusions

In this paper, we have examined a causal model of fertility and marriage, using data from the Sri Lanka Fertility Survey. Although the extreme assumption that there were no net socioeconomic differentials in fertility in Sri Lanka was embedded in the model, it nonetheless performed quite well. Although the model assumption that the entire causal import of socioeconomic factors upon fertility is channelled through age at first marriage and region is undoubtedly a simplification of reality, there is little question that these are major vehicles by which socioeconomic forces work themselves out in cumulative fertility.

The factors examined in this paper do not go very far in determining age at first marriage. A variable which clearly plays such a signal role in fertility merits deeper understanding. Efforts to understand how marriage markets operate and to identify the normative pressures operative on young men and women to marry could be a significant step in this direction.

Although the variables studied in this paper have been investigated by numerous others in various settings, we are not aware of any other effort to treat them in quite the same way as they have been explored herein. Thus, one of the difficulties in interpreting our results has been the lack of benchmarks by which we could judge our results as plausible or implausible or assess whether the estimated effects were relatively large or small. We regard undertaking comparative analysis to put the present results in perspective as first order business.

Finally, we would like to close with a modest caution about the use of path analysis in inquiries of this type. We believe it is useful in helping one sketch out the broad outlines of the interrelationship between variables such as the ones we have considered. In and of itself, however, path analysis cannot yield a full understanding of how these variables are interconnected. Accomplishing that requires resort to additional techniques, including log linear analysis and multiple classification analysis which enable one to examine the data for interaction. Nonetheless, path analysis can

give one a broad outline of main effects from which further more detailed and typically more limited analyses may depart.

#### X. Acknowledgements

We are grateful to the Government of Sri Lanka for permitting us to use the data of the Sri Lankan Fertility Survey conducted in 1975 as part of the World Fertility Survey program. We also greatly appreciate the support and assistance provided by the staff of the Nihon University Population Research Institute.

### Note

1. It should be noted that giving each nominal level variable its best starting place is not equivalent to assigning to it the largest possible effect, net of the other variables, that it could have on fertility. The latter would be accomplished in the following way. Represent all but one of the categories of each nominal level variable by a dummy variable. Then regress the number of children ever born on all of these dummy variables plus the interval and ratio variables. The categories of the nominal variables are then assigned scores equivalent to the coefficients the dummy variables representing them received in the foregoing regression, with the omitted category in each nominal classification being assigned a score of zero. This procedure assigns to each nominal variable its maximum possible net impact upon fertility (where net means independent of the remaining variables --nominal and otherwise--under consideration).

Appendix Table 1. Classifications of Each Nominal Variable Used in Analysis

Variable	Number of Cases	Mean Number of Children Ever Born
Region of Current Residence		
Zone I	634	4.32
Zone II	698	4.45
Zone III	869	5.48
Zone IV	577	5.58
Zone V	566	4.73
Zone VI	1,430	4.91
Type of Childhood Residence		
Urban	950	4.56
Rural	3,413	5.13
Estate	411	4.13
Religious Group		
Buddhist	2,717	5.00
Hindu	1,185	4.74
Christian	396	4.60
Muslim	476	5.50
Racial Group		
Sinhala	2,929	4.91
Sri Lankan Tamil	991	4.89
Indian Tamil	385	4.38
Sri Lankan Moor	469	5.53
Work Status Before Marriage		
No Work	3,399	5.10
Worked	1,375	4.50

Appendix Table 2. Means and Standard Deviations  
of Variables Used in Analysis

Variable Description	Symbol	Mean	Standard Deviation
Type of Childhood Residence <sup>a</sup>	U	4.927	0.3318
Year of Birth	Y	1939.21	7.258
Racial Group <sup>a</sup>	R	4.927	0.2453
Religious Group <sup>a</sup>	C	4.927	0.2240
Years of School Completed	E	4.128	3.365
Work Status Before Marriage <sup>a</sup>	W	4.927	0.2684
Age at First Marriage	H	17.69	3.534
Region of Current Residence <sup>a</sup>	A <sub>1</sub>	4.927	0.4413
Duration of Marriage (in months)	D	207.6	98.76
Number of Children Ever Born	F	4.927	2.805

<sup>a</sup> Nominal variable coded according to mean number of children ever born to women in category.

## References

- Ball, R.J. 1973. The International linkage of National Economic Models. North-Holland Publishing Company.
- Blau, P.M. and O.D. Duncan. 1967. The American Occupational Structure. New York: Wiley.
- Chan, K.C. 1976. "The Role of the Family Planning Association in Hong Kong's Fertility Decline," Studies in Family Planning, 7, October, pp. 284-289.
- Cho, Lee-Jay, and Robert Retherford. 1973. "Comparative Analysis of Recent Fertility Trends in East Asia," paper prepared for the International Population Conference, Liege, Vol. 2, pp. 163-181.
- Fernando, D.F.S. 1976. "Trends and Differentials in Fertility," in Population of Sri Lanka, Country Monograph Series No. 4, United Nations Economic and Social Commission for Asia and the Pacific, Bangkok, pp. 154-172.
- Freedman, Ronald, et al. 1970. "Hong Kong's Fertility Decline, 1961-68," Population Index, 36, pp. 3-17.
- \_\_\_\_\_. 1973. "A Comment on 'Social and Economic Factors in Factors in Hong Kong's Fertility Decline' by Sui-ying Wat and R.W. Hodge," Population Studies, 27, pp. 589-595.
- Goldberger, A.S. 1970. "On Boudon's Method of Linear Causal Analysis," American Sociological Review, 35, pp. 97-101.
- Klein, L.R., and A.S. Goldberger. 1955. An Econometric Model of the United States, 1929-1952. Amsterdam: North-Holland Publishing Company.
- Knodel, John and Nibhon Debavalya. 1978. "Thailand's Reproductive Revolution," International Family Planning Perspectives and Digest, Vol.4, No.2, Summer, pp. 34-49.
- Ogawa, Naohiro, and J.R. Rele. 1981. "Age at Marriage and Cumulative Fertility in Sri Lanka," Asian Population Studies Series No. 49, United Nations Economic and Social Commission for Asia and the Pacific, Bangkok, pp. 227-268. This will be also published as part of WFS Scientific Reports (forthcoming).
- Ohbushi, Hiroshi. 1976. "Demographic Transition in the Process of Japanese Industrialization," in Japanese Industrialization and Its Social Consequences, edited by Hugh Patrick. University of California Press.
- Ryan, Bryce. 1952. "Institutional Factors in Sinhalese Fertility," Milbank Memorial Fund Quarterly, Vol. 20, No. 4, pp. 359-381.
- Sewell, William H. et al. 1970. "The Educational and Early Occupational Status Attainment Process: Replication and Revision," American Sociological Review, 35, December, pp. 1014-1027.

- Tinbergen, J. 1939. Statistical Testing of Business-Theories.  
Volume 1: A Method and Its Application to Investment Activity.  
Volume 2: Business Cycles in the United States of America, 1919-1932. Geneva: League of Nations.
- Wat, Sui-ying and R.W. Hodge. 1972. "Social and Economic Factors in Hong Kong's Fertility Decline," Population Studies, 26, pp. 455-464.
- World Fertility Survey. 1978. World Fertility Survey-Sri Lanka: First Report. Department of Census and Statistics, Ministry of Plan Implementation.
- Wright, S. 1921. "Correlation and Causation," Journal of Agricultural Research, 20, pp. 557-585.