ECONOMIC GROWTH CENTER

YALE UNIVERSITY

Box 1987, Yale Station
New Haven, Connecticut

CENTER DISCUSSION PAPER NO. 382

EVALUATION OF THE OLSEN TECHNIQUE FOR ESTIMATING THE FERTILITY RESPONSE TO CHILD MORTALITY

James Trussell* and Randall Olsen

August 1981

Notes: *Office of Population Research, Princeton University

The authors thank Bryan Boulier and T. Paul Schultz for very helpful suggestions and advice. The work was supported by grant number RF-79036 from the Rockefeller Foundation.

In a recent article, Olsen (1980) proposed a technique for estimating the extent to which child deaths are replaced. The child replacement effect is an important issue in the demography of developing countries. In the 1960's it was argued that efforts to reduce child mortality would induce fertility declines since couples in developing societies produced many children in order to ensure that at least some survived to adulthood. Once it was recognized that a large fraction of children no longer died, couples would adjust fertility accordingly. Therefore, efforts to reduce mortality were to be welcomed not only because of their intrinsic worth but because fertility would fall as well. Such arguments are predicated on the not unreasonable assumption that couples have a certain desired family size defined in terms of surviving children. A replacement effect could come about in two ways. First, actual deaths could elicit a response from the parents to replace directly the lost child. Second, high mortality conditions could create generalized behavior responses which insure that sufficient numbers of children survive. The most obvious example would be hoarding -- creation of a buffer stock of children. They were given more precision by the computer simulations of Heer and Smith (1968). By the end of the decade it had become clear that rapid falls in mortality were not being matched by falls in fertility and that the consequent rapid surge in population growth was a very real problem. It was therefore recognized that some attempt should be made to quantify the impact of reductions in child and infant mortality on fertility. This concern led to research which was presented at a conference held in 1975 sponsored by CICRED. Those

results were summarized by Preston (1975) in an extremely important review article. The conclusion was that the replacement effect was very weak. About 25% of child deaths were replaced in developing countries like Bangladesh, Senegal, and Morocco, but this effect was purely biological, since there was no real control of fertility. In developing countries in which some contraception was practiced (e.g. Colombia, Peru, and Mexico), the replacement effect was even smaller. Even in more developed countries like Taiwan and Costa Rica only about a quarter of deaths were estimated to have been replaced. Taken at face value, these estimates left health programs far from exonerated of the charge that in addition to being socially desirable in their own right, their demographic consequences were deleterious. Rapid population growth was apparently the price which a developing country paid for effective health programs.

More recent work has concentrated on the knotty problem of estimating the extent of replacement. The survey by Schultz (1976) describes much of this literature. The Olsen technique is the culmination of this line of research. He addresses the problem of estimating the direct response of fertility to a child death. This direct effect includes volitional responses of the couple to replace the lost child as well as biological responses via shortened periods of breastfeeding. The impact of fertility hoarding is not included, and hoarding may play an important role in the replacement phenomenon. If so, even if Olsen's estimates are correct, they will understate the importance of replacement. We refer the reader to Olsen's arguments that conventional measures of replacement are biased and to his proposed solution.

Our purpose here is to test the Olsen technique. One problem, heretofore, has been the lack of validation of any technique. Methods have been applied to real data sets. We have no way of knowing, however, whether the results are near to or far from the truth. We avoid this problem by creating a data set for which we know the true answer. To do so, we simulate a set of reproductive histories for which we know the true extent of replacement behavior and then examine the success with which the technique suggested by Olsen estimates replacement. In the next section, we present the model for simulating reproductive histories, describe the simulation, and discuss possible definitions of replacement. Next we briefly review the Olsen technique. Then we test the ability of Olsen's techniques to measure the degree of replacement. In a final section, we summarize our findings.

Monte Carlo Simulations

We employ the simulation model which has been used so effectively by John Barrett (1971). The version which we use is one which we wrote ourselves, but we have followed his suggestions closely. The model is a Monte Carlo micro-simulation in which the reproductive histories of a sample of women are created. The salient features of the model are outlined below.

A. Fecundability is distributed according to a beta distribution with parameters 3 and 9, giving a mean fecundability of .25. Each woman's fecundability declines linearly from age 30 until the end of her reproductive span.

B. The distribution of sterility follows the model specified by Pittinger (1973). The proportion sterile at any age is given by $s(a) = 1.01155 - \exp[-k(r^{a-12}-1)/\ln(r)]$, $12 \leq a \leq 50$. The parameter values $k = .0002$ and $r = 1.251242$ were found to give a nice fit to models proposed by others. These values give model proportions sterile by exact ages 25, 30, 35, 40, 45, and 50 of .027, .060, .154, .389, .778, and 1.0, respectively.

C. Possible pregnancy outcomes are live birth, stillbirth, and fetal death. The probabilities of fetal deaths (γ_2) and stillbirths (γ_3) are age dependent, but do not vary among women of the same age; $\gamma_2 = 0.24 + 0.005(\text{age}-30)$ and $\gamma_3 = 0.03 + 0.001(\text{age}-30)$.

D. Fetal deaths are distributed exponentially from month 1.0 to month 10.0. Fetal losses in month 0 (the first month) are not observable and are absorbed into fecundability. Live and stillbirths have an associated period of pregnancy of 10 months.

E. The period of postpartum insusceptibility is two months for a stillbirth or a fetal death (one month if the fetal death occurs in the second month) and for a live birth the sum of a constant two months and a random variable distributed as a negative binomial with parameters 2 and .1667 (hence the sum has a mean of 12 months). Death of the child truncates the period of postpartum insusceptibility.

F. Age at death of a child is determined from the West model life table, level 17 ($q_1 = .071$, $e_0 = 60$) taken from Coale and Demeny (1966).

G. When used, contraception is 95% effective (fecundability is reduced by 95%).

H. All women start their reproductive careers at age 20 and are observed at age 50. Each simulation is based on 5,000 women.

I. When appropriate, a woman is initially assigned a desired family size of 3 (30%), 4 (40%), or 5 (30%).

We designed five runs which employ different reproductive strategies:

1. Full replacement. Desired family size is framed in terms of surviving children. Contraception is practiced whenever the number of children desired is less than or equal to the number surviving.

2. Mixed replacement and natural fertility. Half the women do not contracept at all; the other half follow Strategy 1.

3. Natural Fertility. No woman contracepts.

4. No replacement. Desired family size is framed in terms of children ever born. Women contracept when the number of children ever born is greater than or equal to desired family size.

5. Mixed no replacement and natural fertility. Half the women do not contracept at all. The other half follow Strategy 4.

Clearly replacement could be modeled in various ways. The one we chose seems to us to be the most natural. Modeling a strategy of no replacement is difficult, because such strategies (except in the case of natural fertility) seem rather artificial. We cannot imagine that couples would actually use contraception in a strategy like case 4. Nevertheless, the definition is clear and the strategy is most easily compared with case 1. Natural fertility is included because others have demonstrated that one can measure a replacement effect which is purely biological. There clearly is no replacement behavior at all in

cases 3, 4, and 5. Mixed strategies are included because it is possible that a technique might correctly identify pure strategies but fail to detect a mixture. Since in real populations a mixture of strategies is most likely to occur, it is crucial that any technique be able to measure the extent of replacement in such situations.

Having designed the experiments so that we know the reproductive strategies, we are still faced with the problem of determining what the technique should be measuring when there is replacement. One approach is to contrast fertility under the alternative strategies of replacement and no-replacement. For example, the mean parity increases from 4.5558 in case 4 to 4.9032 in case 1, a difference of .3474 births per woman. The mean number of deaths in the two cases are .6508 and .7002, respectively. Thus comparing the two situations, .6508 deaths per woman with a strategy of no replacement translates into an increase of .3474 births per woman when the replacement strategy is adopted; thus 53% of deaths are replaced. Actually, however, some of the additional births result from additional deaths. An alternate measure is, therefore, 3474/7002 = 50%, or an average of the two = 51.5%. That this way to measure replacement is the preferred way is by no means certain; it might appear appealing because the only difference between the two simulations is the absence of a replacement motive. A necessary condition for a true replacement motive -- family size framed in terms of surviving children -- is absent in case 4.

This exercise is helpful in sharpening our thinking about how replacement should be measured. There is a severe problem with the

previous measure. Quite obviously, the absence of replacement behavior could be modelled in many ways; these would lead to different measures of the replacement effect if the replacement and no-replacement strategies were contrasted. It is useful to recall that we are interested in assessing the effect of mortality on fertility. Specifically, we seek to infer what would happen to fertility if mortality fell. Hence the straightforward approach is to examine directly by simulation the effect on fertility if mortality is eliminated entirely. With no mortality, the mean parity in case 1 would be 4.5210 instead of 4.9032. Hence, on average, the elimination of one child death causes a reduction of .55 (=(4.9032 - 4.5210)/.7002) of a birth. This way of measuring replacement is the one we adopt. Note that our measure includes both volitional and biological components. No estimator could be expected to separate the two.

The true replacement effect in case 1 is very close to the earlier measure of 50% obtained when cases 1 and 4 were contrasted. This result is no accident, and it strengthens our decision to model the strategy of no-replacement as we did. When there is no mortality, the number of children ever born and surviving is identical. Hence the reproductive strategy in case 4 corresponds to that in case 1 in a situation of no mortality. The actual simulation in case 4 does not correspond exactly, however, to a no mortality situation since this period of post partum insusceptibility may be truncated by a child death. This biological effect contributes an average of .035 births per woman to the total difference of .382.

While this numerical value of .55 is suggestive of the degree of replacement in case 1, it too is subject to sampling error and should not be confused with the 'true' replacement effect which would emerge from an infinitely large simulation. For example, if we calculate the true replacement effect in simulations of size 1000 for case 1 the five true values produced a range of about .15. In the absence of infinitely large simulations it is always possible that apparent error in the estimator could reflect an erroneous notion of what the true replacement effect really is. In order to determine the true population (not sample) replacement rate, we used very large simulations of size 20,000. When the true population rates are calculated, they are found to be 53%, 27%, 10%, 2%, and 7% in cases 1 through 5, respectively.

Note that the effects are not zero in cases 3, 4, and 5 even though there is no replacement strategy because lactation is sometimes terminated by a child death sooner than it would have otherwise been. It is also important to note that even if the replacement effect is fully operative, a decrease in mortality is not matched by an equal decrease in fertility so that population growth increases. The reason for this result is, of course, that even with a full replacement strategy as we defined it, not all couples are successful in making up for deaths. In order to ensure that there is a full making up, couples would have to adopt some hoarding strategy.

We can see this result more clearly, perhaps, by viewing the replacement strategy from the woman's perspective. Although this is not the definition nor the emphasis ordinarily used, let us

concentrate solely on actual replacement, which is defined to occur only when desired family size has been reached. The motivation for this definition is as follows. A woman desires four surviving children. She has one, which dies. But she does not in fact replace this child because she would have continued to try to reach her desired family size even if the child had not died. That is, her reproductive strategy is not altered. So we now concentrate only on women who at one time have achieved their desired family size. They begin to contracept. Some have unwanted births; some have children die. How many must stop contracepting because the number of surviving children falls below the desired number? In case I, 1812 women (36%) stopped contracepting; of these, 1596 (88%) successfully regained their desired family size and began to contracept. Of this number, 171 again experienced a (net) child loss and stopped contracepting, and 113 were successful once again in attaining their desired family size. Of the 113, 13 again dropped below and six successfully regained their desired family size.

From this perspective, then, replacement can occur only when the initial stock is complete. It is clear from this example that not all couples are successful in making up for dead children. The extent of replacement is not fully captured by these figures, of course, because a woman whose first child died and who then went on to bear four children (her desired number), none of whom died before the woman was age 50, would not be counted. Such a situation is, however, reflected in the measure of replacement which we have adopted. Note that our preferred measure of replacement does not necessarily reflect the

extent to which desired family sizes are reached. The reason is that in the alternative state of the world (i.e. no mortality), all other factors are the same. Thus, for example, if contraceptive effectiveness were very low and if desired family size were very low (say 1), almost all couples (except the very infecund) would achieve at least the desired family size at the expense of large numbers of unwanted births. Still, under our way of measuring the replacement effect, fertility would fall by less than mortality if mortality were reduced to zero. Hence, since some couples would be unable fully to replace dead children, the replacement effect under our replacement strategy would always be less than one. To get an effect of at least one, the additional strategy of hoarding must be adopted.

By using the same mortality schedule for all children, the only variation in the fraction of a couple's children that die will be due to differential exposure of children born at different dates and the luck of the draw. This assumption completely rules out heterogeneity due to differences in family specific nutrition, exposure to disease and living conditions. It is rather unlikely in practice that such variation does not exist. To generate such heterogeneity, the Coale Demeny schedule of proportions dead by age x was multiplied by four times the woman's fecundability. Since mean fecundability is .25, the mean proportionality factor is 1.0.[1] This modelling scheme assumes that $q_x = 1 - 1_x$ schedules are proportional; such an assumption has been used extensively and has empirical grounding (see Trussell and

---

[1] The factor of proportionality is below 0.4 (producing an e_o of about 70) in 10% of the cases and above 1.7 (producing an e_o of about 50) in another 10% of the cases.

Preston, 1981). Simulations for cases 1 through 5 were repeated under this assumption of correlated fertility and mortality.

This very strong positive correlation between mortality and fertility greatly biases the naive regression estimates and is a stringent test of Olsen's methods. In addition, a positive correlation could result from fertility hoarding strategies, so this set of simulations may also be viewed as a test of the ability of Olsen's methods to distinguish pure replacement of dead children from the related strategy of replacing anticipated deaths. The true population replacement rates in this second set of simulations are very close to those presented earlier: 53%, 26%, 10%, 3%, and 8% in cases 1 through 5, respectively.

The Olsen Technique

A brief summary of the methodology is now in order. There are two regression estimators upon which the final estimates of replacement (denoted $\tilde{r}$) are based.

1) The ordinary least squares (OLS) estimates, denoted by r_{OLS} which is obtained by regressing births n_i on deaths d_i;
2) the instrumental variation (IV) estimates, denoted r_{IV} which is obtained in a two-step process. First d_i is regressed on the proportion dead $p_i = d_i/n_i$. The predicted values of d_i from this regression ($\tilde{d}_i$) are then employed as the regressors; n_i is regressed on $\tilde{d}_i$ (not d_i as in (1) above).

All regressions contain a constant term. The unit of observation is the woman (family); women with no births are excluded altogether because they can provide no information on the relation between fertility and mortality.

The OLS coefficient is always a biased and inconsistent estimate of the true replacement rate. However, Olsen developed correction factors, described below, for a variety of different circumstances. The IV estimate is sometimes consistent; under some circumstances, however, it, too, must be corrected. How does one know which correction factor is needed?

The main diagnostic tool is the implied within parity variance of the mortality rate $\sigma^2_{p|n}$. This statistic is implied because the mortality rate is unobservable. It is estimated as

$$\sigma^2_{p|n} = [\text{var}(d_i|n) - n\bar{p}(1-\bar{p})]/(n^2-n) \qquad (1)$$

for each value of n, where $\bar{p}$ is the average proportion of dead children in the entire sample and $var(d_i/n)$ is the sample variance of the number of dead children per woman of parity n. Since this statistic is estimated, it can take on (impossible) non-positive values. These are included when the average value across parities is computed; to exclude them would introduce a systematic bias. Before describing the diagnostics and correction factors in more detail, we must take a brief detour to extend the methodology slightly.

Mixed strategies are included in our simulations because it is likely (indeed, virtually certain) that there exists a mixture of strategies in real populations. Since Olsen did not consider mixed strategies it was first necessary to examine the statistical theory in order to determine which estimator will be preferred in such cases. When couples can follow different strategies, the problem can be viewed as a random coefficients model. The analysis is confined to Appendix 1 in order that the flow of our presentation not be unduly interrupted. The conclusions reached are that the OLS-based estimate of replacement is likely to be biased downward but that it is possible to construct a consistent estimate for r by correcting the instrumental variables (IV) estimator:

$$\tilde{r} = r_{IV} - \rho\sigma_p\sigma_n/cov(d_i/n_i, d_i) \qquad (2)$$

where r_{IV} is the instrumental variables estimator, $cov(d_i/n_i, d_i)$ and σ_n are estimated using their sample moments, and σ_p (the standard deviation of the mortality rate) and ρ (the correlation between the mortality rate and fertility) are estimated using the method suggested

by Olsen (see Appendix 2). The latter methods for estimating ρ and σ_p will be robust to the random coefficient problem since they rely upon the sample means and variances of n and d and the mean of p and its within parity variation. It is of no consequence to that method whether part of the variability in n is due to a random replacement coefficient since the behavioral replacement equation is not used in these derivations.

We offer the following rules as a guide for selecting which particular estimator is appropriate:

A) If the observed variance of d_i in the sample is very close to

$$\bar{n}\bar{p}(1-\bar{p}) + \bar{p}^2 \text{ Var}(n) \tag{3}$$

and the implied within parity variances are close to zero (standard deviations on the order of .01) or negative, then this is an indication that across all women the probability of a child death is constant. The method for a nonstochastic mortality rate is then appropriate; that is the OLS estimate is adjusted as

$$\tilde{r} = r_{OLS} - \left[(1-\bar{p}r)[\bar{p}+(1-\bar{p})\bar{n}/\text{var } (n)]\right]^{-1} \tag{4}$$

where $\bar{n}$ and var(n) are the sample mean and variance of children ever born and $\bar{p}$ is the average mortality rate (total deaths/total births) in the sample. Note that r appears on the left and right sides of equation (4). One must solve for r, therefore, by an iterative procedure. We suggest that the

investigator start with a guess of r and produce a new estimate by using equation (4). This process continues until convergence is achieved. IV, with no correction, may also be used in this case to provide a consistent estimate of r.

B) If the observed variance of d_i in the sample is very close to

$$\bar{n}\bar{p}(1-\bar{p}) + \bar{p}^2 \mathrm{Var}(n) + \mathrm{Var}(p|n)[\mathrm{Var}(n) + \bar{n}^2 - \bar{n}], \qquad (5)$$

where Var(p|n) is the average implied within parity variance of the mortality rate, then the mortality rate can be taken as random but uncorrelated with fertility. Instrumental variables (IV) with d_i/n_i as the instrument can be used to obtain consistent estimators of r; no correction is needed to the IV estimate. An alternative is to use OLS with the following corrections:

$$\tilde{r} = r_{OLS} - \bar{p}\ \mathrm{Var}(n)/\mathrm{Var}(d) \qquad (6)$$

C) If the average implied within parity variance in mortality rates is positive but Var(d) is not well approximated by (5), then there is evidence that the mortality rate is random and correlated with fertility. In this case the nonlinear equations in Olsen must be solved, preferably for both a bivariate lognormal distribution for n and p and a normal-lognormal distribution for n and p. This procedure yields a correction to be applied to the least squares coefficient. The estimates for σ_p and ρ which are also obtained using

this method can then be used to correct the IV estimator according to the formula given in equation (2). A step by step procedure is given in Appendix 2.

C') If, in case C above, there is a number of negative values of the implied within parity variance of the mortality rate, then caution must be exercised. Such values may be symptomatic of a more fundamental specification error.

D) If the implied average within parity variance in mortality rates is very small or negative, indicating no within parity variation in mortality rates, and Var(d) is different from its predicted value in (3), the methods in (B) and (C) cannot be applied, and the model implicit in (A) will be misspecified. In this instance, the confidence attached to the estimates should not be unduly high. Because the method in (A) tends to be a lower bound estimate, a better choice would be to use $\tilde{r} = r_{OLS} - \bar{p}\ \mathrm{Var}(n)/\mathrm{Var}(d)$. This is essentially the same estimator as in (4) except that Var(d) is used in the place of $\bar{n}\bar{p}(1-\bar{p}) + \bar{p}^2\mathrm{Var}(n)$ and $\bar{p}r$ is taken to be very small. Instrumental variables may also be used in this case, although it will not be possible to diagnose or correct the problems which arise when fertility is correlated with mortality. The closer together are the two estimates, the higher should be the confidence that they are capturing the true replacement effect. If $\bar{n}\bar{p}(1-\bar{p}) + \bar{p}^2\mathrm{Var}(n)$ is larger than Var(d) it is very likely that fertility and

mortality are negatively correlated and the corrected OLS method and uncorrected IV will tend to underestimate true replacement. If the inequality is reversed, little can be said, although ceteris paribus, a larger Var(d) implies these methods are more likely to overstate replacement.

E) A finding that the corrected instrumental variables estimate is higher than the corrected OLS estimates may be a sign of random coefficients. In such an event the IV estimate (corrected for a correlation between fertility and mortality if necessary) is the preferred estimate. It is difficult to know how much higher than the OLS based estimate the IV based estimate must be in order that it be preferred. We have adopted a rule of thumb of 50% higher; otherwise the average of the two is chosen.

In summary, when p_i and n_i are uncorrelated we can use the IV estimator or correct our least squares estimator. If p_i and n_i are correlated we can correct either the instrumental variables coefficient or the least squares coefficient. When the two methods give different results, with the instrumental variables estimate being substantially higher, the discrepancy may be due to random coefficients, in which case the instrumental variables based estimator would be preferred.

Results

We first consider the case where the children of all women face identical mortality schedules. A constant problem with these simulations was that the implied within parity variances in the mortality rate were most often negative. This property puts the Olsen methods at a disadvantage since the implied within parity variation in p_i plays an important role in diagnosing the stochastic structure of the data. In such cases both the OLS and IV estimates are subject to great uncertainty. If the IV estimate is much higher, it is preferred; otherwise an average of the two is preferred.

In Table 1 some of the samples were subdivided to investigate the effect of sample size. The five smaller samples for cases 1 and 2 came from the first large sample for cases 1 and 2, respectively. Both methods did well for case 1, but the investigator should approach the results warily since the stochastic structure could not be diagnosed. This result may just be a quirk in the simulation; when examining the Colombian data Olsen never found negative implied within parity variances in the motality rates, so in practice D diagnoses may be quite rare. Note that the range of values for $\tilde{r}$ for the small samples of case 1 using the OLS correction is close to range of true sample replacement values using the smaller samples.

We turn next to the mixed strategies embodied in case 2. As the statistical theory predicts, the least squares correction method tends to underestimate replacement when the coefficients are random. The IV estimates in the two case 2 samples of size 5000 were very close to the true population replacement effects. For all the case 2 samples

of size 1000 in Table 1, the average of the IV estimates was too low by .09 and the variability was very high. Finally, in cases 3, 4, and 5 the preferred estimates captured nicely the minor biological replacement effects.

When the mortality schedule was made to be random across couples the results in Table 2 were obtained. In case 1, IV overstates the extent of replacement, which is to be expected since there was no way to solve for the correlation between fertility and mortality. The corrected least squares estimate overstated replacement slightly. In case 2 the presence of within parity variation in the mortality rate enabled IV to estimate replacement accurately. The large difference between the instrumental variables and the corrected OLS coefficient correctly suggested random coefficients. In cases 3, 4, and 5 both methods tended to capture the biological replacement effect. In case 3, IV was sufficiently high to raise suspecions of random coefficients.

The results presented thus far pertain to women who are observed at age 50. Investigators, however, are most often interested in determining the effect of mortality on fertility among younger women without having to wait until they reach age 50. There is keen interest in assessing whether replacement behavior is currently changing by examining differences among cohorts. A final set of simulations was designed to test whether the Olsen procedure could be used on a sample of women who had not reached the end of their reproductive careers. To do so, we truncated the observation period for the simulations in Table 1 at ages 40 and 30. Results are shown

in Table 3. Two points are worth noting. First, the true replacement rate, measured from the contrast between the mortality and no-mortality situation, varies as the age at observation changes. This result is hardly surprising, since women are caught at different points in their reproductive life cycle. These differences, however, are relatively small. Second, the Olsen technique produces estimates of replacement which are rather close to the true values at all three ages at observation. The discrepancy at the younger ages is largest for cases IV and V. The evidence from the simulations suggests that the technique can be used on younger women, though perhaps with not as much confidence as on women whose reproductive careers are complete. A finding of large and systematic differences across cohorts would suggest that replacement behavior is changing.

Summary

The simulations here are not meant to assess completely the sampling properties of the estimators; only a full Monte Carlo experiment can do that. Nevertheless, there is some indication that sample sizes as large as a thousand may still result in estimates with considerable variability. In some ways the simulation data did not replicate important features of real data, such as substantial within parity variation in mortality rates, which greatly hindered the application of Olsen's methods. When the data allowed random mortality to be handled satisfactorily, (those uses diagnosed as B, C, or C' in Tables 1 and 2) estimates within .07 of the true values were obtained in ten of eleven cases. In the other cases, the stochastic structure could not be diagnosed; the average error was -.05, and the variability of the errors was quite large.

We have examined the sensitivity of these results to changes in assumptions employed in the simulations. Trial calculations show that the results are robust to the choice of mortality schedule (level and shape), age at marriage (including incorporating a distribution of age at marriage), and level of contraceptive effectiveness.

Our summary evaluation is that the technique performs well, especially in cases where the stochastic structure of the data can be diagnosed. This finding bolsters Olsen's previous result that there is evidence of a replacement effect in Colombia. If the Colombian data Olsen used are reexamined, the average direct replacement rate is 0.24; this figure is higher than the 0.18 originally reported because the technique did not allow for the possibility of random

coefficients. If half the correlation between fertility and mortality were due to hoarding, the total replacement rate in Colombia would be in the vicinity of three quarters (Olsen, 1980). We are currently engaged in a project to determine the magnitude of the effect in other developing countries. One final point is worth repeating. As our simulations show, the measured replacement effect understates considerably the proportion of the population who employ a replacement strategy. The difference between the fertility effect and the proportion adopting the strategy is of course due to the stochastic nature of the process; some may not need to replace and others may not be successful even if they do try.

Table 1: Fixed Mortality Schedule

Case	Observations	Diagnosis	IV-Based Estimates	OLS-Based Estimates	Average[$]	True Population Rate
1	5000	D	.46	.46	.46*	.53
1	1000	D	.56	.53	.54*	.53
1	1000	D	.37	.40	.38*	.53
1	1000	D	.42	.39	.40*	.53
1	1000	D	.48	.46	.47*	.53
1	1000	D	.47	.40	.43*	.53
2	5000	D,E	.24*	.12	—	.27
2	1000	C',E	.20*	.01	—	.27
2	1000	D,E	.46*	.24	—	.27
2	1000	D,E	.23*	.11	—	.27
2	1000	D	-.04	-.02	-.03*	.27
2	1000	C',E	.06*	-.03	—	.27
2+	5000	D,E	.30*	.15	—	.27
3	5000	B	.04	.04	.04*	.10
4	5000	B	.01	.02	.01*	.02
5	5000	C'	.07*	.04	—	.07

Notes:

* - indicates preferred method based upon diagnosis

A - nonrandom mortality rate

B - random mortality rate - uncorrelated with fertility

C - random mortality rate - correlated with fertility

C'- same as C except that negative within parity variances may degrade analysis even though average within parity variance is positive.

D - Random mortality rate but negative average within parity variance makes further diagnosis impossible

E - indication of random coefficients

\+ - case 2 samples were generated by a random selection of half the observations in case 1 and half in case 3. Hence there were two samples of size 5000 available for analysis.

$ - when the IV-based estimate is not clearly preferred, the final estimate is the average of the two.

Table 2: Random Mortality Schedule

Case	Observations	Diagnosis	IV-Based Estimates	OLS-Based Estimates	Average[$]	True Population Rate
1	5000	D	.66	.61	.64*	.53
2+	5000	C',E	.21*	.05	—	.26
2+	5000	C',E	.22*	.06	—	.26
3	5000	C,E	.09*	.06	—	.10
4	5000	C	.03	.03	.03*	.03
5+	5000	C	.09	.09	.09*	.08
5+	5000	C'	.13	.11	.12*	.08

Notes:

* – indicates preferred method based upon diagnosis
A – nonrandom mortality rate
B – random mortality rate – uncorrelated with fertility
C – random mortality rate – correlated with fertility
C'– same as C except that negative within parity variances may degrade analysis even though average within parity variance is positive.
D – Random mortality rate but negative average within parity variance makes further diagnosis impossible
E – indication of random coefficients
\+ – case 2 samples were generated by a random selection of half the observations in case 1 and half in case 3. Hence there were two samples of size 5000 available for analysis.
$ – when the IV-based estimate is not clearly preferred, the final estimate is the average of the two.

Table 3: Estimates of replacement effect when observation is truncated at ages 40 and 30

	Age at Observation					
	30		40		50	
Case	Estimate	True	Estimate	True	Estimate	True
I	.46	.54	.61	.62	.45	.53
II	.26	.33	.36	.31	.24 (.30)	.27
III	.07	.14	.03	.11	.04	.10
IV	-.03	.06	-.04	.01	.01	.02
V	.00	.11	-.08	.07	.05	.07

REFERENCES

Barrett, John (1971). Use of a fertility simulation model to refine measurement techniques. Demography 8: 481-490.

Coale, Ansley J. and Paul Demeny (1966). Regional Model Life Tables and Stable Populations. Princeton University Press: Princeton.

Heer, David and Dean Smith (1968). Mortality level, desired family size, and population increase. Demography 5: 104-121.

Olsen, Randall J. (1980). Estimating the effect of child mortality on the number of births. Demography 17: 429-444.

Pittinger, Donald (1973). An exponential model of female sterility. Demography 10: 113-121.

Preston, Samuel (1975). Health programs and population growth. Population and Development Review 1: 189-200.

Schultz, T. Paul (1976). Interrelationships between mortality and fertility in R.G. Ridker (ed.), Population and Development. The Johns Hopkins University Press: Baltimore.

APPENDIX 1

Estimates with Mixed Strategies

Let us modify the behavioral model which Olsen uses in the following way:

$$n_i = \bar{n} + r_i^*(d_i - \bar{d}) + u_i \qquad \text{(A1)}$$

where our use of r_i^* indicates that each couple i follows a different replacement strategy where $r_i^* = \bar{r} + \bar{\eta}_i$. This is a random coefficients model and can be reexpressed as

$$n_i = \bar{n} + \bar{r}(d_i - \bar{d}) + u_i + \eta_i(d_i - \bar{d}) \qquad \text{(A2)}$$

where the regression coefficient $\bar{r}$ is the average replacement rate in the population. Because the random replacement coefficient introduces a new error component into the residual, $\eta_i(d_i - \bar{d})$, Olsen's expressions for plim($\tilde{r}$) for least squares are no longer valid. It is necessary to evaluate the mathematical expectation $E[d_i - \bar{d})^2 \eta_i]$ in order to derive these probability limits. Note that d_i and η_i are not independent since <u>ceteris paribus</u> a larger η_i (more replacement) implies more deaths since the couple will have more births as it seeks to replace. The evaluation of $E[(d_i - d)^2 \eta_i]$ requires information on the joint distribution of $(d_i - \bar{d})$ and η_i. If these two random variables were bivariate normal, then this expectation would be zero. This situation would tend to lessen the bias of the least squares estimate of r since the covariance of d_i and u_i would be lessened to the extent that part of the variance in n_i could be attributed to variation in η_i.[2] The estimated replacement

rates based upon the derivations of plim ($\tilde{r}_{OLS}$) will tend to be too low so long as $E[(d_i-\bar{d})^2\eta_i]$ is small relative to $E[(d_i-\bar{d})u_i]$, as seems likely.

If η_i is independent of p_i, then d_i/n_i will be uncorrelated with $u_i + \eta_i(d_i-\bar{d})$ so that instrumental variables using d_i/n_i as an instrument will be consistent, provided that p_i is uncorrelated with u_i.

If p_i is independent of η_i but correlated with u_i, then since $d_i/n_i = p_i +$ terms uncorrelated with u_i, the probability limit of the instrumental variable estimator is $r + \text{cov}(p_i,u_i)/\text{cov}(d_i/n_i,d_i)$. By solving for the correlation of p_i and u_i, the variance of p_i,[3] and taking $\text{Var}(u_i) = \text{Var}(n_i)$ as an approximation for small r, it is possible to construct an estimator for r by correcting the instrumental variables coefficient. The result is shown in equation (2) in the text.

---

[2] Note that we assume that $\bar{r}^2\text{Var}(d)$ is a small fraction of Var (n).

[3] Note that $\bar{p}_i$ represents the probability that a child of couple i will die before the mother reaches 50. Variation in $\bar{p}_i$ across couples can reflect different exposures to death or different mortality tables. Variation in p_i within families is of no consequence to the estimators.

APPENDIX 2

Estimation when Fertility and Mortality Rates are Correlated

When the mortality rate is correlated with fertility, higher order moments of the joint distribution of fertility and the mortality rate must be evaluated in order to obtain the proper correction for the least squares regression coefficient. If ln(n) and ln(p) follow a bivariate normal distribution where the mean of ln(n) is μ_x, the mean of ln(p) is μ_y, the variance of ln(n) is σ_x^2, the variance ln(p) is σ_y^2, and the correlation of ln(n) and ln(p) is ρ, then

$$E(n^r p^s) = \exp[r\mu_x + s\mu_y + (r^2\sigma_x^2 + s^2\sigma_y^2 + 2\rho rs\sigma_x\sigma_y)/2] \qquad (B1)$$

Using this formula we can solve for expressions involving sample moments, for example

$$\mathrm{Var}(n)/E(n)^2 = \exp(\sigma_x^2)-1 \qquad (B2)$$

so

$$\log\left[[\mathrm{Var}(n)/E(n)^2] + 1\right] = \sigma_x^2 \qquad (B3)$$

where we can use $\bar{n}$ in the place of E(n) and the sample variance of n for Var(n). Now ρ is the latent correlation between ln(n) and ln(p), so

$$\mathrm{Var}[\ln(p)|\ln(n)] = \sigma_y^2\,(1-\rho^2). \qquad (B4)$$

Just as the marginal variance of ln(n) could be expressed as a function of the mean and variance of n, we can do the same for the conditional variance of p given n, so

$$\log[1 + \mathrm{Var}(p|n)/E(p)^2] = \sigma_y^2(1-\rho^2). \qquad (B5)$$

Once we pick a value of ρ we can solve for σ_y, and, together with σ_x, produce all the other required moments. That is,

$$\mathrm{Var}(p) = \bar{p}^2\left[\exp[1/(1-\rho^2)](1 + \mathrm{Var}(p|n)/\bar{p}^2)\right] \tag{B6}$$

$$E(np) = \bar{n}\bar{p}\exp(\rho\sigma_x\sigma_y) \tag{B7}$$

$$E(np^2) = \bar{n}(\mathrm{Var}(p) + \bar{p}^2)\exp(2\rho\sigma_x\sigma_y) \tag{B8}$$

$$E(n^2p^2) = (\mathrm{Var}(n) + \bar{n}^2)(\mathrm{Var}(p) + \bar{p}^2)\exp(4\rho\sigma_x\sigma_y) \tag{B9}$$

$$E(n^2p) = (\mathrm{Var}(n) + \bar{n}^2)\bar{p}\exp(2\rho\sigma_x\sigma_y) \tag{B10}$$

and the correlation between n and p is

$$\frac{E(np)-E(n)E(p)}{[\mathrm{Var}(n)\mathrm{Var}(p)]^{1/2}} = \frac{\exp(\rho\sigma_x\sigma_y)-1}{[(\exp(\sigma_x^2)-1)(\exp(\sigma_y^2)-1)]^{1/2}} \tag{B11}$$

From the appendix of Olsen

$$\mathrm{Var}(d) = E(np) - E(np^2) + E(n^2p^2) - E(np)^2 \tag{B12}$$

and the bias of the least squares regression coefficient is

$$[E(pn^2) - \bar{n}E(pn)]/\mathrm{Var}(d). \tag{B13}$$

Various values of ρ are selected until the value of Var(d) as computed above is suitably close to the sample variance of d. We then use this value of ρ to estimate the correlation between n and p and the bias of the least squares estimator.

If we assume n and ln(p) have a bivariate normal distribution we can apply a very similar method. Again we use

$$\log[1 + \mathrm{Var}(p|n)/E(p)^2] = \sigma_y^2(1-\rho^2) \tag{B14}$$

which yields a value of σ_y for given ρ, where ρ is the correlation between n and ln(p). We set $\sigma_n^2 = \text{Var}(n)$, $E(p) = \bar{p}$ and $\text{Var}(p) = \bar{p}^2(\exp(\sigma_y^2)-1)$.

The required higher order moments are now

$$E(np) = (\bar{n} + \rho\sigma_n\sigma_y)\,\bar{p} \tag{B15}$$

$$E(np^2) = (\bar{n} + \rho\sigma_n\sigma_y)(\text{Var}(p) + \bar{p}^2) \tag{B16}$$

$$E(n^2p^2) = (\text{Var}(p) + \bar{p}^2)(\sigma_n^2 + (\bar{n} + 2\rho\sigma_n\sigma_y)^2) \tag{B17}$$

$$E(n^2p) = \bar{p}(\sigma_n^2 + (\bar{n} + \rho\sigma_n\sigma_y)^2). \tag{B18}$$

The same procedure is followed; pick a value of ρ which equates the sample variance of d to the function of the above population moments and then obtain the bias of the least squares coefficient and the correlation of p and n ($=[E(np) - E(n)E(p)]/[\text{Var}(n)\text{Var}(p)]^{1/2}$).