A REVIEW OF RECENT DEVELOPMENTS IN THE ECONOMICS OF FERTILITY

by

Diane J. Macunovich

June, 1993

RP-164

GIANNINI FOUNDATION OF AGRICULTURAL ECONOMICS

WILLIAMS COLLEGE
WILLIAMSTOWN, MASSACHUSETTS
A REVIEW OF RECENT DEVELOPMENTS IN THE ECONOMICS OF FERTILITY

by

Diañe J. Macunovich

June, 1993

© 1993-(Diane J. Macunovich)

Note: This paper is intended for private circulation and should not be quoted or referred to in publication without the permission of the author.
A Review of Recent Developments
In the Economics of Fertility

Diane J. Macunovich

Department of Economics
Williams College
Williamstown, Massachusetts

January 1991

Abstract: This paper attempts to review the literature on the economics of fertility and to synthesize the findings of that literature in order to see how they fit in the context of the New Home Economics (NHE) and the Easterlin relative income models. Emphasis is placed on the need to work with an exogenous measure of the female wage, and on the possibility that the income effect of the female wage may be stronger than has been assumed in NHE models. Results of tests of both models are presented and discussed, and reasons are suggested for the ‘failure’ of the two models to explain fertility in the 1980s. A call is made for further attempts to combine the two models, and potential problems in doing so are discussed.

This paper is a draft of a study to be included in the forthcoming book Household and Family Economics, edited by Paul Menchik and published by Kluwer Academic Press. Its preparation was supported in part by NICHD Project Grant No. P-50-HD-12639.
Following its heart-stopping roller coaster ride in the decades after World War II, the U.S. total fertility rate settled into the doldrums, hovering around 1.8 for the better part of the 1970s and 1980s. The economy rose and fell, taking with it wages and employment rates; the baby boom gradually assumed its place in the labor force as the focus shifted to the baby bust generation; and female labor force participation rates continued to climb — but the fertility rate stolidly refused to budge significantly in either direction. Why bother trying to model something that isn’t moving? In the United States the ‘official response’ to fertility patterns over the past two decades is summed up well in the following statement from Patrick Thompson et al (1989):

"...demographers contend that the large swings in fertility during the baby boom and bust of 1945-1974 were generated by a fertility process that is not likely to be repeated; the stability in fertility rates since 1974 has been cited to support this argument. Consequently, in the Census Bureau projections the high and low fertility variants [of TFR] were constrained to 2.2 and 1.5, respectively."

But in the last three years the U.S. has witnessed its first sustained rise in the total fertility rate in three decades. It appears that the 1990 rate may have climbed back up to 2.1, the replacement level (see Figure 1). Is this merely a temporary aberration? Even if it is, is there an explanation for it? Are we any further along in understanding the economic underpinnings of the fertility rate than we were in the 1970s? We certainly have more data — longer time series, and more comprehensive micro data, which have generated increasingly sophisticated econometric analyses.

What do these data indicate about the validity of our economic models? This paper is an attempt to answer that question. It is an attempt to step back from the detailed and highly technical nature of many of the studies which have been pursued in the last decade — to synthesize some of our findings in order to see how they fit in the context of the Easterlin and New Home Economics models.

Both the New Home Economics and the Easterlin models are based on the common assumption of an underlying positive relationship between income and fertility, and both attempt to put forward explanations for the negative relationship which has often been observed (both in time series and in cross section) in modern experience. Both in turn attempt to explain the postwar baby boom and bust. They differ fundamentally, however, in their identification of the driving force behind these movements, as indicated by the labels used to describe them: the ‘relative income’ model and the ‘value of time’ model (Schultz, 1978).

More significantly, perhaps, these two models differ in terms of their prognostications for the future. If the Easterlin model is correct, then the recent upturn in fertility rates is probably a harbinger of another baby boom — perhaps a series of recurrent booms — and recent estimates (Ahlburg & Vaupel, 1990) indicate that, if such a pattern were to occur, it would more than double the population of the U.S. by the year 2080. Even a ‘moderate’ boom (with fertility peaking at only 2.7, rather than at 3.2) would
add another 176 million people to the population by 2080.\(^1\) As Ahlburg and Vaupel state, such a fluctuation would have a markedly more dramatic effect on U.S. population levels and age structure than any other set of combined demographic changes considered possible within the foreseeable future.

If, on the other hand, the New Home Economics model best describes our current situation, then the recent upturn is simply an unexplained 'hiccup,' to use Peter Morrison's phrase (Pendleton, 1990), and we can expect to see further declines in the TFR as more and more women follow the lure of increasing female wages, and put career before the demands of motherhood. In this case, we face a longterm trend of below replacement fertility, with all its attendant implications regarding the aging of the population and resultant strains on the Social Security system.

So it seems a pertinent time to review these secular models. The attempt of this paper will be to demonstrate that even the past two decades of fertility behavior are explainable within the context of a relative income model, once appropriate controls for female wages and labor force participation have been introduced, and the asymmetric nature of cohort size effects has been recognized. If this explanation has any validity, then the recent upturn in fertility rates is the result of rising relative male wages in combination with an increasingly dominant income effect of female wages.

---

\(^1\) Both of these rates used by Ahlburg and Vaupel are well below the 'post-war' peak of 3.77, reached in 1957. In the 1970s and 1980s the TFR hovered around the 1.8 level.
Sections A and B describe the New Home Economics and Easterlin models briefly, and then discuss issues involved in attempting to combine the two models, while Sections C and D describe the results of recent tests of these two models. Section E then reviews recent findings in the literature on female labor supply, regarding the effects of the female wage, and Section F examines the implications of these findings, for models of fertility. Sections G through I look at why the models have, or have not, been predicting fertility trends accurately, and Section J then assesses implications for future fertility trends.

A. A Brief Overview of the Contenders

The New Home Economics Model. The New Home Economics model suggests that parental demand for children can be treated as analogous to the demand for producer or consumer durables, depending on whether parents expect net pecuniary returns from children or receive direct utility from them. In this model, the cost of children is endogenous, since parents receive utility from increased child ‘quality’ as well as from increased numbers of children. As a result, an increase in total expenditures on children caused by increased family income might be devoted largely to increased expenditures per child rather than to increases in the number of children. In addition, the model makes the cost of children endogenous because of the price of time expended on childrearing activities. This model was originally suggested by Becker (1960), and then elaborated by Becker and Lewis (1973). Mincer (1963) first suggested that the wife’s wage could be used as a measure in this model, of the price of time in childrearing.

Willis (1973) elaborated Mincer’s suggestion and demonstrated that this negative price effect could offset a positive family income effect, thus producing a negative relationship between income and fertility even if parents did not desire to increase the standard of living of their children as their own standard of living increased. His attempt to test this model using micro-level data was not successful, but his model has been used as the basis for a number of studies using macro-level data, most notably the time series analysis of Butz and Ward (1977, 1979a). In Willis’ model, and in these macro-level formulations, the emphasis has been on tests of the price and income effect (with the wife’s wage as a measure of price, and the husband’s earnings as a measure of income), rather than on the trade-off between child quantity and quality.

In this New Home Economics model it is important to note that the model can be interpreted as strictly causal only if increases in female wages are exogenous. That is, to the extent that the female wage, and hence the price of time in raising children, is simply a function of increasing levels of female education and/or experience, we must continue our search for, as Schultz (1986, p. 98) put it, ‘an initial cause, or the forcing developmental variables’. Without the demonstration of exogeneity, the New Home Economics Model incorporates ‘child quality/quantity tradeoff’ effects and effects of intergenerational endowments, as well as the price of time effect, the focus in this paper is on the price of time effect, since this is the only aspect of the model which has been put forward as an explanation of the postwar baby boom and bust.
Economics model is simply an explanation of the observed negative relationship between income and fertility, using a neoclassical Slutsky decomposition into income and substitution effects.

Butz and Ward developed a model which represents one extreme of the Becker-Mincer-Willis model, and they used this model to attempt to explain the 'baby boom and bust' and associated short-term fluctuations in U.S. fertility. In this model, it is postulated that both medium and short-term fluctuations can be attributed to a dominant price effect of a rising female wage (which they treat as exogenous), and subsequently varying levels of female labor force participation. This model posits that the baby boom is attributable to rapidly increasing male earnings at a time when the female participation rate, and thus the aggregate negative price effect of female wages, was low. The baby bust, on the other hand, occurred as a result of rapidly increasing female wages which attracted increasing numbers of women into the labor force and at the same time increased the opportunity cost of the wife's time. With increasing numbers of women in the labor force, the negative price effect of increasing female wages began to dominate any positive income effect of male incomes. The driving force in this model is thus an assumed exogenous increase in female wages.

The Easterlin Model. Whereas the New Home Economics model endogenizes the cost of children, the Easterlin model endogenizes the preference for children and uses relative income as an indicator of this endogenous shift in preferences (Easterlin, 1966). Easterlin postulates a systematic shift in preferences resulting from the fact that each successive generation, under economic development, experiences a successively higher parental standard of living. "In effect, a... 'subsistence level' constraint is added to the analysis of [fertility behavior] along with the budget line and production constraints." (Easterlin, 1978a; and Ahlburg, 1984). Because of this 'subsistence level' constraint, economic or demographic fluctuations could cause periodic reversals in the secular downtrend in fertility, such as that observed in the developed countries in the postwar period. In the full Easterlin model, relative cohort size inversely affects cohort earning potential, which in turn directly affects cohort fertility and female labor force participation.

According to the Easterlin theory, a large birth cohort meets unfavorable labor market conditions which reduce the earning potential of young males relative to their aspirations. In an attempt to close the gap between income and aspirations, members of such a cohort will tend to make a number of adjustments including increased female labor force participation and delayed/reduced marriage and childbearing. Thus in this formulation the driving force behind both increased female labor force participation and reduced fertility, is the desire of a large cohort to improve relative economic status, with parental income as the measure of that cohort's material aspirations.

B. Can the Two Models be Combined?

There have been a number of calls to combine these two models: for example, Paul Schultz stated in 1978 that

"as one approach appears to complement the other, an integrated approach promises to be even more useful, reducing bias attributable to
omitted variables and to explanatory variables that are probably endogenous in the more comprehensive framework." (Schultz, 1978, p.27).

—and similar arguments have been made by Warren Sanderson (1976, 1980). Easterlin provided qualified support for the inclusion of the female wage as an indicator of the price of time in raising children (See, for example, Easterlin 1978a, pp. 65-66). Superficially this seems eminently feasible: a framework in which relative income, rather than absolute income, is expected to exert a positive influence on fertility, while a combination of income and substitution effects operate through the wife’s wage.

However, it is important to bear in mind in any attempt to combine the two models, that there is a very fundamental difference between them. In particular, the two models postulate very different mechanisms leading to increased levels of female labor force participation. In the New Home Economics (NHE) Model, exogenous increases in female wages increased the opportunity cost of women’s time in the home, and led to increased female labor force participation and decreased fertility. The NHE model is causal to the extent that increases in the female wage are demonstrated to have been exogenous — and simply descriptive otherwise. Thus, in the NHE model it is entirely appropriate to include the unadjusted female wage — if the intention is to be purely descriptive — as a variable representing the cost of time in raising children. However, use of the New Home Economics model to define causal relationships—as in Butz and Ward (1977, 1979a)— depends heavily on the demonstration that increases in the female wage were indeed exogenous. This problem is acknowledged in Schultz (1978, p.33).

In Easterlin’s model, however, the primary causes of these two phenomena (decreasing fertility and increasing female labor force participation) was the fall in young males’ earning potential relative to their aspirations. There is no mention in this model of any exogenous change in the female wage. Thus in this case it would be appropriate to include the female wage only to the extent that increases in the wage are exogenous, or to the extent that we include controls for education and work experience, treating both of these latter variables as functions of relative income. This treatment would recognize that to a large extent observed increases in the female wage result from increasing levels of female work experience, tenure, and possibly human capital investment resulting from the need to participate more fully in the labor market.

It is, perhaps, one of the most significant shortcomings of Easterlin’s model, that the actual mechanics and implications of the rise in female labor force participation were never fully elaborated. Subsequent work on female labor supply does, however, provide us with considerable material for bridging that gap in the model.

Why is this difference so important? In the case of the New Home Economics model, because the female wage is the driving force, its function as a price variable in a fertility equation, and the absence of a labor force participation variable in that same equation, are unquestioned. As Mincer wrote,

"The economic analyses of fertility and of labor-force participation here presented suggest that the choices of labor and family size are not
causally related to one another. Rather, these choices are simultaneously determined by the same basic economic variables. The higher the female wage rate and the lower the husbands’ earning power, the higher the labor-force rate and the smaller the fertility rate. The relation between fertility rate and labor-force participation is not autonomous; it does not provide new insight once the two structural relations are specified. And indeed, as the empirical findings indicate, a labor-force variable \( X_4 \) introduced in addition to the income and price variables is redundant; it adds little or nothing to the explanation. (Mincer, 1963).

However, in Easterlin’s model the precise role of the female wage and the female labor force participation variable in a fertility equation have not been elaborated. As a result, because of the many calls to ‘combine’ the relative income and price of time models, the standard approach (Butz and Ward 1977, 1979a; Ermisch 1979, 1980) has been simply to graft the major components of the two approaches: convert male earnings to relative male earnings as a measure of the positive income effect, and add in a female wage variable as the price of time. This hasty and ill-thought out approach, however, neglects the fact that in order to support causal interpretations of the model, only exogenous increases in the female wage can be used as independent variables. It must also be recognized that in the Easterlin model, the female wage must necessarily represent a much stronger income effect than is postulated in the New Home Economics Model. This is not to deny that it would also exert a price effect, but rather to emphasize that if, as Easterlin postulated, women entered the labor force in order to bring households up to a minimum consumption threshold below which children were considered an untenable option, then the wife’s contribution to household income played a central role as income in determining whether or not that household could afford to support a child.

If Easterlin’s model has any relevance as a causal model, then there is a substantial causality from relative male income to female participation, so that there is a large endogenous component in the female wage. Female wages increase at the individual level as a result of increasing levels of experience (on the job human capital accumulation), and also to some extent as a result of increased incentive to invest in education given the expectation of greater labor market participation. At the aggregate level these changes might be positive or negative depending, to use Mincer’s terminology (1985) on whether ‘at the margin, increases are extensive or intensive.’ That is, if the female population prior to the 1960s was heterogeneous, (composed largely of two types of women: those who worked regularly outside the home and those who worked not at all outside the home) then increases in the size of the female labor force would tend to bring in, for the most part, women with little aggregate labor force experience who would thus tend, initially to bring down the overall average wage. If, on the other hand, women were

\[^3\text{Mincer's } X_4 \text{ was the aggregate age-specific female labor force participation rate. Mincer has since acknowledged, however, that female labor force participation does have a place in a fertility equation. (Columbia Labor Economics Workshop, November 1989).}\]
homogeneous (all spending some time working in the home and some time in the labor force), then an increase in labor force participation would mean an increase, at the margin, in aggregate labor force experience and hence in average wages.

C. Recent Tests of the Easterlin Relative Income Model

What does the literature of the past decade tell us about the relevance of the relative income hypothesis? We will first examine a series of tests at the aggregate level, and then move on to micro-level analyses.

1. Relative Income Effects at the Aggregate Level

Macro-level tests of the Easterlin relative income model using U.S. data have generally supported the hypothesis, while initial micro-level applications such as that of MacDonald and Rindfuss (1978) have not. Examples of supporting macro-level studies are those of Moffitt (1982) Devaney (1983) and Schapiro (1988). In addition, Butz and Ward (1977, 1979a) tested two variations of the relative income variable, in conjunction with male income and female wages, and found a "significant positive coefficient as hypothesized," but concluded that the role of the relative income variable was "not a dominant or consistent one" (p.326). However, Devaney notes that "by ignoring relative income as a predictor of female labor supply, Butz and Ward fail to identify an indirect effect of relative income on fertility" (p. 148).

Devaney addressed this shortcoming of the Butz and Ward study by estimating a set of simultaneous equations for age specific fertility and female labor force participation, using an auxiliary female wage regression and U.S. aggregate data for the period 1947-1977, provided by Butz and Ward. She found a significant positive effect of relative income on fertility during the reproductive years and a significant negative effect on female labor force participation at all ages.

In Europe, however, tests by Ermisch using German (1980) and British (1979, 1988) data, and tests by Wright (1989) using data from 16 Western European countries found only weak support for the Easterlin hypothesis, leading to Ermisch's (1990) conclusion that "Easterlin's... relative income theory has also been used to explain the baby boom and subsequent baby bust as well as changes in women's employment, marriage and divorce. I have not discussed it here because, while there is support for it from U.S. experience, the evidence does not suggest that it plays a major role in European fertility changes."

However, two of Ermisch's tests of Easterlin's model using European data rely on an unusual formulation of the relative income variable. It is a formulation due to Wachter (1975), which uses the ratio of two time series of earnings for all males, undifferentiated by age. That is, Ermisch (1979, pp.43-44) uses "the time series of 'expected men's real wages.' This series is obtained by deflating the money wage by the retail price index and adjusting it by multiplying it by the probability of employment (equal to the complement of the male unemployment rate). The time series of 'expected male real
wages' \( (W) \) is then used to construct a series of the ratio of 'son's wages' to a hypothetical 'father's wage'. . . The 'son's wage' in year \( t \) is, therefore, defined as the average of \( W \) for the seven previous years and the 'father's wage' is defined as the average of \( Wt \) for the ten years preceding \( t-5 \). The ratio is the 'relative wage. . .'

while Ermisch (1980, p.126) uses:

"two series for the desired standard of living by calculating average men's expected real wages for the previous seven and ten years for each year \( t \). The ratio of men's expected real wages in year \( t \) to this desired standard of living in that year is denoted as the relative wage."

That is, both the numerator and the denominator in these relative wage series reflect earnings of males of all ages — the only difference between the numerator and the denominator being the seven or ten year lag of the latter.

This is a crucial shortcoming in his tests, since the mechanism proposed by Easterlin (and later substantiated by Welch, 1979; Berger, 1984; and Macunovich and Lillard, 1989) for reductions in relative income rests on the imperfect substitutability between younger and older workers in the labor force, and therefore the inverse effects of a change in cohort size on earnings of the two groups. An increase in the size of the younger cohort will tend to decrease the earnings of the younger group and increase the earnings of the older group.

Also, as indicated by Macunovich and Lillard (1989), not only earnings are affected by relative cohort size: employment rates, participation rates and hours and weeks worked are all adversely impacted by increases in relative cohort size, so that Ermisch's use of wages as opposed to earnings, and his correction for employment expectations using the employment probability of males regardless of age, also severely understated the effects of relative cohort size.

"Since hours worked by the husband and the extent of the wife's gainful employment are both a function of the husband's real wage the latter is a better measure of the resources influencing family decisions than either actual family income or husband's earnings, both of which are endogenous variables resulting from household choice." (Ermisch, 1980 p. 126 and 1979 p.43).

Thus by using a time series of earnings for all men, Ermisch missed these crucial effects of cohort size on earnings — and may also for this reason have missed the effect of relative income on fertility.

Turning to Wright's (1989) examination of the relative income hypothesis, we find an even less satisfactory measure of relative income: gross measures of relative cohort size. He takes up the effort of Chesnais (1983), who attempted to relate changes in the net reproductive rate to relative cohort size using 'a visual inspection of the trends and simple zero-order correlations', which Wright found too simplistic (Wright, 1989, p. 109):
"One crude way of testing the Easterlin hypothesis is to construct a measure of ‘relative cohort size’, which Easterlin considers to be a suitable proxy for relative income. This measure is usually expressed as a simple ratio of the male population aged 30-64 to the male population aged 15-29 (i.e., N_{30-64}/N_{15-29}). Upward movement in this ratio indicates a decrease in the number of ‘young’ people relative to ‘older’ people, which is assumed to correspond to an increase in relative income for the younger group. It follows that changes in relative cohort size should exhibit a direct positive association with changes in fertility." (Wright, 1989, p. 109).

Ermisch, too, in his 1988 study used a relative cohort size measure to proxy relative income effects. In this case, he used ‘relative generation size of the cohort [of women] born in [the relevant time period] (relative to the 1948 cohort)’ (Ermisch, 1988, p.564).

The use of these measures was forced on Wright and Ermisch by the difficulties inherent in constructing age-specific income series. Wright does acknowledge problems with the measure in the statement ‘However, given the undoubtedly crude way in which [relative cohort size] “measures” the central theoretical concept of relative income, the Easterlin hypothesis may still be valid.’ (Wright, 1989, p.113). However, the use of these crude measures highlights an issue which is central to tests of the Easterlin hypothesis: the fact that there has been little, if any, research on the formation of material aspirations, so that the formulation of a relative income measure even at the micro level is fraught with hazards -- while at the macro level there is so little adequate guidance that one might almost toss a coin in attempting to decide between alternative measures.

At the micro level, do we need a comparison of earnings of the husband relative to his father, to her father, or both? And should these be comparisons of wages or earnings or total income? What is the relevant period during which the father’s earnings can be said to have affected his child’s material aspirations? Do we want earnings at a specific age, or expected lifetime earnings? And should they be corrected for family size? Not only have none of these questions been answered to date: in general we lack micro datasets which provide information on first and second generation income and fertility behavior (but see the discussion of Behrman and Taubman (1989 and 1990) and Macunovich and Lillard (in progress) below).

Thus tests of the hypothesis have generally been carried out at the macro level — but this leads to innumerable difficulties in the specification of an adequate proxy for relative income. Which age groups should be used in the numerator and denominator — and what lags, if any, should be used in the denominator? These decisions are crucial to the definition of the shape and turning-points of a relative income measure.

For example, in Wright (1989) the numerator was defined as males aged 15-29, but if males aged 20-29 had been used instead, the turning points of his relative income measure would have been made to correspond much more closely to turning points in fertility. It is arguable that, given the later ages
at marriage typical in Europe relative to the U.S., the use of the 20-29 age group would have been more appropriate. Ermisch (1988) indicates an awareness of this sensitivity of results to choice of numerator and denominator, in his discussion of a relative wage measure he devised for females:

"For instance, these measures of net earnings produce a ratio of women’s net hourly earnings to men's net weekly earnings (RW) that rises from the mid-1960s until a peak in 1976, after which it falls. In light of the fact that conditional birth rates generally peak in the mid-1960s, reach a trough around 1977 and then recover, it is not surprising that the coefficients of RW consistently are negative and significant." (Ermisch, 1988, p. 571).

In this case, Ermisch is justifying additions to and deductions from gross earnings of females, which differ from those used in his earlier studies — rather than the choice of age group. But the point remains, that in defining aggregate measures of what is intrinsically a microlevel phenomenon, a certain amount of arbitrariness and subjectivity is almost unavoidable.

Perhaps even more critical in Wright’s 1989 study is the fact that cohort size is assumed to be an adequate proxy of relative income — however relative income might be defined — at the aggregate level. However, Easterlin has stressed in all of his work on long swings that he saw relative cohort size effects operational (in the sense that they might be expected to produce relative income shifts) only in economies which are relatively closed to migration. Thus, in the U.S. prior to the mid 1920s he does not claim that relative cohort size shifts in the indigenous population would have led to relative income shifts: instead, migration flows would have acted to equilibrate supply and demand side effects. (See, for example, Easterlin, 1987, pp. 32-33; and 1968, pp. 13).

This is a consideration which has been overlooked to a large extent in tests of the Easterlin hypothesis in the European context. To the extent that the EEC has caused a number of countries to be integrated as an economic unit, and has encouraged flows of labor across international borders, relative income shifts would become increasingly dissociated from indigenous relative cohort size shifts. In addition, the relatively free use of ‘temporary’ workers in European countries over the past two decades (and to some extent, perhaps, the flows of illegal immigrants into the U.S.) would have had a depressant effect on the earnings of young, unskilled workers4 — just at the time when their relative cohort size in the indigenous population was becoming more favorable. For these reasons, it is highly doubtful that a relative cohort size measure could be expected adequately to proxy relative income over the past few decades in Europe.

Also, it is important to note that none of the studies decisively rejected the relative income

---

4 See, for example Johnson (1980), Wachter (1980), Houstón (1983) and Chiswick (1989). Also, Macunovich and Lillard (1989) found that geographically disaggregated studies which indicate small or negligible effects of immigrants on low-skilled natives might be biased due to the tendency of immigrants to be attracted primarily to areas of high wages and low unemployment.
hypothesis, despite their poor proxies:

"These results suggest that though relative economic status may not be the dominant influence upon fertility suggested by Easterlin, it may yet be important among a number of factors influencing fertility." (Ermisch, 1979, p. 49)

"Empirical analysis also suggests that movements in real income relative to experience-based material aspirations have played a role in postwar fertility movements, albeit a secondary one." (Ermisch, 1980, p.140)

"... some support was found, however, in the post-1950 fertility trends in Belgium, England and Wales, France, Finland and Italy. ... The evidence presented above is far from being conclusive. The model and tests carried out are dependent on the assumption that the index of relative cohort size is a 'good' if not perfect, proxy for the underlying theoretical concept of relative income. As mentioned earlier, some empirical evidence suggests this may not be the case in countries other than the United States. (See Ermisch, 1979 and 1980). In light of this, we therefore must strongly qualify our overall conclusion: the empirical analysis indicates that the hypothesis that relative income is the dominant influence on fertility must be rejected on the basis of the experience of the European countries considered here, if the selected index of relative cohort size is taken as a good measure of relative income. It is clear that research aimed at examining the magnitude of the correlation between relative cohort size and relative income is needed before firmer conclusions concerning the relevance of the Easterlin hypothesis outside the American context can be made." (Wright, 1989, p.118).

"The simulation 'tracks' births well, and counter-factual experiments suggest that changes in women's pay relative to men's, in real house prices and in relative generational size are primarily responsible for the fluctuations in births during 1971-85." (Ermisch, 1988, p.575)

2. Relative Income Effects at the Individual Level

Tests of Easterlin's hypothesis at this level have suffered historically from the paucity of data containing socioeconomic information on two related generations of individuals. Thus, studies which were conducted in the 1970s, such as MacDonald and Rindfuss (1978), relied on measures such as responses to the question 'How well-off are you?' or on comparisons of education or occupation levels of fathers and sons, to proxy something approaching Easterlin’s measure. These were clearly inadequate, and we are only now beginning to exploit the few sources of data which permit some matching of parents and their adult children. One of these is the University of Michigan’s Panel Study on Income Dynamics (PSID), which is the source of data for a relative income study currently in progress by Macunovich and Lillard (supported by NICHD Project Grant No. P-50-HD-12639).
Another is the NAS-NRC Twin and Adult Offspring Sample used by Behrman and Taubman (1989, 1990), in combination with an offspring survey they carried out themselves in the early 1970s. These datasets are limited, in that the parent portion contains records only for white male twins both of whom served in the military — and at the time these twins’ offspring were surveyed they had an average age of only 27, and only about one-half were married, so that total expected rather than actual number of children was used in Behrman and Taubman’s study. In addition, Behrman and Taubman limited their sample to offspring who were currently married and whose spouses were in the labor force.

Behrman and Taubman in these two studies attempt to distinguish Easterlin’s relative income model from Becker’s model of intergenerational serial correlation in endowments (which Becker says produce indistinguishable results), and in doing so to test which of the two models is supported by the data. Their first attempt is a ‘special case’ of the first, and compares the coefficient on a variable measuring parental income and with another indicating number of siblings: Their model is:

\[ n_t = b + \eta I_t - \pi n_{t-1} + \pi n_{t-1} + x_t \]

where \( n_t \) is the completed family size of generation \( t \), \( I_t \) is the income in generation \( t \), and the coefficients \( \pi \) on \( I_t \) and \( n_{t-1} \) are expected to be equal in magnitude but opposite in sign, on the assumption that the relevant income measure for the parental generation is a per capita one.

Behrman and Taubman state that their results here do not support the Easterlin model, since they reject the null of equality of the absolute values of these two coefficients. However, this is a very restrictive test, since it assumes that the number of siblings enters only in terms of the per capita income calculation. If, however, there is some positive (negative) serial correlation between generations in preferences with regard to desired family size (which would bias upward (downward) the coefficient on \( n_{t-1} \)), then their null hypothesis does not represent an adequate test of Easterlin’s model.

Perhaps in recognition of this limitation on their first test, Behrman and Taubman (1990) then attempted a more general test using a latent variable approach incorporating information from individuals, cousins and siblings to control for unobserved intergenerational serial correlations such as that alluded to in the previous paragraph (in the same dataset described above). Here, their results

"are statistically significant and have opposite signs in the two equations. The sign pattern is that predicted by Easterlin and contrasts with that predicted by Becker... Therefore, our test suggests that the Easterlin intergenerational taste effect is more consistent with recent United States’ fertility experience than is the Becker intergenerational endowments formulation. This contrasts with our results in Behrman and Taubman (1989) but the model used in that paper is a special case of the model developed in this paper, so the present results supersede the earlier ones." (Behrman and Taubman, 1990, pp. 28-30).

---

5 See Manski, 1990, who indicates that ‘intentions data bound but do not identify the probability that a person will behave in a given way’. This is particularly relevant in a relative income analysis, where Easterlin would maintain that while the desired family size has not varied significantly over time, the shortfall between income and aspirations has caused realized family size to vary significantly.
Their work can only be taken as limited support of the model, however, because of the data limitations outlined above, and the highly restrictive form of the model tested (in which total family income was used rather than male earnings for each generation). Perhaps as a result of their formulation of the income variables, the effect of own generation income on fertility was found to be insignificant.

Another microlevel analysis carried out on an Israeli dataset is reported in Danziger and Neuman (1989). Here, although data were lacking on the earnings of the husband’s and wife’s parents, proxies in the form of occupational status were used and found to exert a significant negative influence on second generation fertility. In addition, this study included a predicted wage variable for the wife, which was included along with a control for the wife’s education. While her education was found to exert a strong negative influence on fertility, her wage displayed a strong positive effect, which the authors interpret as indicating a dominant income effect. Husband’s predicted wage was insignificant, however, after controlling for parents’ income (but strongly negative before including such controls). Danziger and Neuman’s results accord very well with earlier findings by Ben-Porath (1973, 1975), also using Israeli data.

Another set of studies can also be read as support for the Easterlin model. They are descriptive analyses, one at the aggregate level (Easterlin, 1987, chapter 10) and the others at the microlevel (Macunovich and Easterlin, 1990; Easterlin et al, 1991a, 1991b). These studies as a group analyze life cycle demographic adjustments made by the baby boomers (such as postponed/reduced childbearing, increased levels of female labor force participation, and doubling up), and find that the net result of these adjustments was a significant improvement in ex post relative to ex ante per capita income.

Finally, no discussion of tests of the Easterlin model is complete without mention of the voluminous work of mathematical modelers, particularly Ron Lee and Ken Wachter, who have attempted to determine whether, given reasonable limits on the various parameters, sustained cycles of the Easterlin type are mathematically possible. One recent paper (Wachter and Lee, 1989) calls into question earlier findings, notably those of Lee (1974) and Frauenthal and Swick (1983), that ‘limit cycle oscillations have been occurring in U.S. births.’

"Though we found that estimates from U.S. data made the model capable of generating sustained cycles, we found the period of sustained cycles to be much longer than Lee had thought or than a fit to U.S. series could tolerate, upwards of ninety years. Transient cycles remained a possibility, but for sustained cycles both the favored special models failed." (Wachter, 1988, p. 2)

However, Wachter (1988) qualifies these findings, by introducing a reinforcing feedback mechanism (the equivalent of a ‘bandwagon’ effect on social norms and attitudes):

"Relative size effects like those emphasized by Easterlin tend to improve models driven primarily by the response of cohort fertility to own cohort size. But they do not on their own improve models enough to match the targets. The most promising further alternative is to introduce a strong
element of feedback into the timing as well as the level of fertility in couples' lives. The full exploration of timing models is beyond the scope of this paper. But the best of the relative size models... come close enough to matching the targets that fine tuning within the constraints already identified holds out promise of success." (Wachter, 1988, p.22)

D. Recent Tests of the New Home Economics Model

The neoclassical, or New Home Economics, model has been tested several times using U.S. and non-U.S. data. It is important to bear in mind in reviewing the results of these tests that the female wage used is normally unadjusted for education or experience, so that while supporting results do indicate a negative correlation between the female wage and fertility, this cannot necessarily be interpreted as indicating a causal relationship, without further controls to assure the exogeneity of the female wage.

1. Price and Income Effects Using Aggregate Data

In non-U.S. applications, Ogawa and Mason (1986) using Japanese data found support for the Butz-Ward model. Subsequently, however, Lee and Gan (1989) found that the female wage had an ambiguous effect on fertility in a 3SLS system which recognized the interrelated nature of fertility, female labor force participation and marital status. This ambiguity arose in large part because of an estimated negative own wage elasticity of female labor supply, in their equations.

Ermisch (1979, 1988) applied the model to British data. In both studies he was unable to obtain age specific wages or participation rates for women, so in the earlier study he proxied the age specific participation rate of married women using the aggregate female employment ratio, and age specific female wages using 'expected men's real wages' (age 21 and over). In his later study, he applied a dynamic model, testing for cointegration between variables, using net after tax female hourly earnings relative to the same for males. However, while his dependent variable is birth order, year and age specific, his income and wage variables are aggregated time series for men and women of all ages. Although Ermisch found a negative effect of the female wage and a positive effect of male income in both of these British studies, his choice of empirical proxies for these variables raises a question regarding the interpretation of the results, over and above the qualification regarding potential endogeneity.

Devaney (1983) conducted a study in which both the relative income and new home economics models were tested using U.S. data. Her findings with respect to the effects of relative income on fertility and female labor force participation support Easterlin's model, but she also found ‘not only that female wage rates are significant predictors of fertility and female labor supply, but also that they are the dominant factor in explaining recent variations in fertility and female labor force participation trends.’ However, the female wage series used by Devaney was that used by Butz and Ward in their studies (1977, 1979a), and has been found by Macunovich (1989,1990) to be seriously flawed, as will be discussed below. Their series indicated a strong increase in the 1960s which is not supported by other data, and continuing increases in the 1970s when other sources indicate a real decline. Perhaps as a result
of these anomalies in the series used, Devaney found a positive effect of relative income on the female wage, which does not accord with the Easterlin model.

A third application by Ermisch (1980) used German data, which he split into two samples, for 1935-77 (excluding 1944-57) and 1957-77. For the later sample he was able to obtain real women’s wages, whereas for the early sample he proxied with an average real wage (all ages 21 and over). In these models, he finds an insignificant effect of the female wage (which he interprets as a canceling out of income and substitution effects), a positive effect of male income in one-earner households and a negative effect of male income in two-earner households (which he interprets as a dominant substitution effect for men). Although these findings appear to be almost diametrically opposed to the hypotheses underlying the New Home Economics model, Ermisch interprets them as supporting that model. He alludes to the fact that Butz and Ward’s own data produce similar results when corrections are made for serial correlation in the data.

Sprague (1988), in an interesting application using British age-specific data for the period 1954-1984, includes three measures for women: average hourly earnings (manufacturing), an education proxy (number of 17 year olds remaining in school as a proportion of all 17 year olds each year), and a permanent wage proxy. Although none of these data used as explanatory variables were age-specific, she obtained relatively significant effects and found that while education and current hourly earnings affect fertility negatively, the permanent wage variable enters with a significant positive coefficient. She interprets this as indicating that ‘Increased permanent wages allow women to be able to afford to have more children, whereas, variations in current wages seem to determine when.’ (p.691). She found a significantly positive effect of male income on fertility.

Smith and Ward (1984, 1985) demonstrate that the average level of experience, and the average wage, in the female population (as opposed to the female labor force) have been rising relative to males in the U.S., and that these increasing average female wages were positively associated with female labor force participation and negatively associated with fertility in the period from 1950 to 1980:

"The dependent variable. . . is the first difference within cohorts in age-specific birthrates. The parameter estimates are conventional, showing large negative elasticities with respect to female wages and smaller elasticities with respect to male wages." (p.87)

Having established the effect of female wages on fertility, Smith and Ward then estimate labor supply equations including both the female wage and predicted fertility as regressors. (It should be pointed out that this approach — including fertility as a regressor in female labor supply equations — has been questioned by a substantial number of researchers, who maintain that both fertility and female labor supply are affected by common unobservables, so that estimated effects of fertility on female labor supply will be biased downward). They summarize their findings:

"With all fertility-induced effects included, real-wage growth 'explains' 58% of the postwar increase in female labor supply." (p.89)
However, the measures of aggregate female wages developed by Smith and Ward, while age and cohort specific, were not education or experience specific. Thus the increases which they calculated were the result of rising levels of education and experience, which disproportionately favored women's over men's wages. To the extent that education is endogenous — that is, to the extent that it is a function of expected monetary returns, or to the extent that it is simultaneously determined along with fertility and female labor force propensities by unobservables — these calculations and regressions of Smith and Ward beg the original question. Similarly, since they hypothesize that female labor force participation is a function of the wage, then experience is also endogenous, and its effects must be eliminated before using an aggregate female wage to explain extensive increases in female labor force participation (that is, entry into the labor force of women with little or no previous labor force experience).

In a simple extension of earlier bivariate analyses initiated by Thomas (1927), Macunovich and Easterlin (1988) attempted a test of the relationship between unemployment rates and fertility in the U.S. using Granger-Sims causality tests, with the objective of determining whether the relationship was procyclical, as indicated by previous researchers, or countercyclical as hypothesized by Butz and Ward. The findings of that study indicated a continuation and even a strengthening of the procyclical fertility observed by Thomas and others.

However, Mocan (1990) challenges this finding by introducing additional variables into the test, and examining impulse response functions. His work verifies the bivariate results established by Macunovich and Easterlin (1988), but indicates that these results are reversed when additional controls are introduced. In particular, he adds variables representing the 'proportion of young marriages' and the divorce rate. His findings indicate that, while the net effect of economic conditions on fertility is still procyclical, this net effect is largely the result of indirect effects operating through age at marriage and divorce. This conflicts, however, with the micro-level findings in Macunovich and Lillard, as reported in Macunovich (1989), where female unemployment was observed to exert a significant negative effect on the hazard of a conception in the first marital birth interval.

The work of Mocan (1990) is very much in the spirit of work conducted by McDonald using Australian data. McDonald (1983), like a number of other researchers, pointed out that serial correlation was a severe problem in Butz and Ward's data, and demonstrated that correcting for serial correlation typically reversed the signs on coefficients and/or removed their significance. McDonald feels that time series vector autoregressive (VAR) models exhibit much better predictive power than the traditional aggregate neoclassical model, once the latter has been corrected for serial correlation.

2. Using Micro-Level Data to Test the New Home Economics Model

Heckman and Walker (1989) took up what they saw as a 'challenge' in McDonald's work, to produce an aggregative neoclassical model which performed better than VAR models that assign no role to wages and income. They estimated a hazard model formulation of a standard neoclassical model of fertility, by birth order and cohort, for individuals sampled in the Swedish Fertility Survey in 1981, which they then aggregated to produce time series trends. The results of their micro-level estimation
indicated what they termed ‘highly nonlinear’ effects of income and wages on fertility — the effects varied by birth order and by cohort of mothers — and they attributed to this nonlinearity the poor performance of traditional aggregate neoclassical models. However, despite nonlinearities, the effect of their female wage variable is always negative and highly significant, and the effect of the male income variable is always positive and generally significant.

It is important to note that the income figures used in Heckman and Walker were aggregate age- and sex-specific averages derived from annual tax data and assigned to the individuals, and that the female wage was estimated by dividing average income thus obtained, by average hours worked. The problems indicated above in Butz and Ward’s estimates of the female wage give cause to question the use of estimates here. Moreover, the female wage figures used by Heckman and Walker were not adjusted for education or experience. As indicated above, while this model may find a significant negative correlation between the female wage and fertility, this cannot be interpreted as a causal relationship without evidence that increases in the female wage were truly exogenous.

Heckman and Walker do address this problem, stating that the Swedish labor market during the study period (1947-81) ‘was characterized by centralized wage setting and explicit egalitarian wage policies’. They assert that this ‘implies that the observed dramatic change in relative female wages from 65% to 90% of male wages... is due to an exogenous force.’ (p. 960) Given the length of the time period analyzed, however, and the changes in education level which occurred, truly exogenous increases in the female wage could only be identified by controlling at least for education, if not for experience as well.

Moreover, even in these regressions which do not adjust the female wage for education and experience, the negative effect of the female wage attenuates noticeably with later cohorts. Heckman and Walker say that ‘Because market work entitles women in the later birth cohorts to larger child benefits, the estimated female wage coefficient reflects less of a substitution effect and more of an income effect on fertility for later cohorts of women.’ (p.960)

Schultz (1978, 1980) alluded to problems of potential endogeneity in attempting to forecast female labor supply, and addressed them directly in later analyses (1985, 1986). He used a predicted wage ‘that does not include as explanatory variables endogenous fertility, family composition, or labor market experience.’ As he explained,

"A woman's realized wage is clearly a function of her past accumulation of labor market experience, and consequently it tends to be inversely related to her fertility and directly related to her life cycle investments in labor market skills. The realized wage rate is, therefore, endogenous to the couple's reproductive choice or plans, or, in other words, the unexplained error in the fertility equation is likely to be correlated with the wife's past, current and anticipated labor market and human capital investment behavior."

However, in his 1986 paper using microdata from the 1967 U.S. Survey of Economic
Opportunity, Schultz did use educational attainment as the primary explanatory variable in his equation for predicting the female wage, despite the problems inherent in this approach. His results indicate a strong negative effect of female wages and positive effect of male income on fertility. However, Schultz acknowledges in that work the endogeneity of the unadjusted female wage used, and therefore the inappropriateness of claims that the female wage was a prime mover in observed fertility changes:

"In many high-income countries the wage received by women relative to that received by men has increased in the last 100 years... Economists hypothesize that this aspect of modern economic growth increases the opportunity cost of children and may be a contributing factor in the secular decline of fertility. But this trend in female to male wages may itself be caused by women's employment and human capital choices, which have evolved simultaneously with the reduction in reproductive goals."

He ultimately concludes that the underlying cause of increasing levels of female education -- and hence of increasing female wages -- is a higher income elasticity of the demand for women's education, than for men's, and a greater responsiveness of female education to changes in the relative cost of higher education.

"The income elasticity of female enrollment is twice as large as that for males, .51 versus .26, and female enrollments are also more responsive to [teacher] prices than are male enrollments, -.92 versus -.69. Evidence in this cross-section of countries suggests that rising real incomes and decreasing prices for schoolteachers have paralleled the advance of women's educational status relative to men's."

This finding is strongly questioned by Lee (1986), however:

"I was surprised at the treatment of secular changes in education here; it appears that education is treated as a sort of consumer good with a high income-elasticity whose demand is therefore driven by rising per capita income. I expected to see it treated here, as it has often been elsewhere, by others and by Schultz himself, as an investment decision in response to rising returns to education. The returns to education rise and stay high because of continuing technical progress, itself perhaps resulting from increased education. Thus the incentive to provide education for oneself or for one's children is that it increases the value of time, which is the same reason that it leads to subsequent reductions in fertility. Viewed in these terms, the ultimate cause of the secular rise in education and the value of time, on the one hand, and secular decline in fertility on the other, remains unclear."
In addition, Schultz's findings of higher income elasticity for female education, as determined from an international cross-section, needs more explanation in the U.S. context, where the average level of education in the male population has increased more rapidly than in the female population, in most of the post WWII era.

3. A Re-examination of Time Series Results Using Aggregated Micro-data

Macunovich (1989, 1990) used microdata from the Public Use Tapes of the March Current Population Survey (CPS), and aggregated these data in order to evaluate the estimated time series prepared by Butz and Ward (1977, 1979a), and to re-estimate their model. Using the CPS, she was able to obtain age-specific wage and labor force participation data for women and annual income for their spouses, and to aggregate these individual records to obtain time series which better reflected the underlying trends, than Butz and Ward's estimates. She was also able to obtain marital fertility rates in addition to the total fertility rates used by Butz and Ward.

When the new Macunovich CPS data series are compared with the Butz-Ward series, several highly significant discrepancies emerge, primarily having to do with their female hourly wage series. A comparison of the CPS and Butz-Ward wage series, together with fertility rates for women aged 20-24 and 25-34, is presented in Figure 2. There, it can be seen that the CPS series provides little support for the notion that female wages increased dramatically throughout the 1960s and 1970s. In Macunovich (1989, 1990), it is demonstrated that the discrepancies between the CPS and Butz-Ward series are attributable primarily to two factors:

- Butz and Ward's use of total annual income of all women, earned and unearned, as a proxy for the annual earnings of working women;
- their use of the average hours worked per week by all workers regardless of sex, in the retail trade, as a proxy for the average weekly hours of all women.

Macunovich found that when the Butz and Ward model was tested using CPS data, the female wage variable was insignificant in its effect on fertility in both age groups, and male income was insignificant on fertility in the 20-24 age group and significantly negative in its effect on fertility in the 25-34 age group. Using marital fertility in place of total fertility as the dependent variable in each age group was found to worsen still further the fit of the model, so that R²s dropped to levels as low as .48.

Even using Butz and Ward's own data, Macunovich found that their model produced insignificant and/or perverse effects for the entire period after 1955 — which is the period in which Butz and Ward felt that the female wage exerted its strongest effects. She found, in addition, that when Butz and Ward's coefficients based on their 1947-1975 data, and updates of their data to 1983, were used to predict fertility in the 1976-1983 period, their model underpredicted actual fertility in that relatively short period by 24% for women aged 25-34 and by 30.8% for women aged 20-24.

These results reported in Macunovich's work might be explained by Heckman and Walker's (1989) finding of significant non-linearities in the data — across birth orders and cohorts. If such non-linearities are present, then models fit on aggregated data will tend to display insignificant or even
Figure 2: Comparison of Age-Specific Fertility Rates and CPS and Butz-Ward Female Hourly Wages ($1964).
perverse results, especially when purged of their serial correlation. In addition, these results must be viewed in the light of findings of Smith and Ward (1984), Goldin (1989), and Heckman (1979) regarding heterogeneity in the female workforce, and the resulting negative effect on average female wages of the entry of large numbers of women with relatively low levels of experience. It is possible that the time series of expected wages in the female population as a whole might have followed a very different pattern than the time series of the wages of working women. Unfortunately, however, Smith and Ward do not provide evidence of absolute increases in the expected wages of the female population — only of increases relative to males. This is less surprising given the decline in real earnings of males in the period since 1973, which would itself have exerted a negative influence on fertility, if the neoclassical model holds. And, in any event, it must be emphasized once again that a correlation between fertility and women’s wages unadjusted for education and experience does not necessarily indicate a causal relationship, nor does it necessarily indicate a negative effect of the female wage per se.

E. Wage Elasticities of Female Labor Supply

One of the unfortunate results of the increasing tendency toward specialization in economics is the inability of researchers in one area to keep abreast of all the recent developments in other areas, in order to be able to incorporate the results of that work into their own. One instance of this information breakdown appears to be the failure of fertility economists thus far to respond to a quiet revolution which has been occurring among economists working in the field of female labor supply.

While the fertility rate stagnated — and perhaps because of it — research on secular trends in western fertility rates became increasingly rare. The bulk of research began to focus instead on life cycle fertility and female labor supply, attempting to calibrate labor supply functions which appeared to be many times more complex — in their estimation, if not in their form — than those of males. It seems that a strong motivating force behind these studies was the emphasis on the male-female wage gap, and attempts to explain that gap on the basis of female work experience and discontinuities due to childbearing. Thus fertility became a side issue: an explanatory or simultaneously determined variable in studies of life cycle female labor force participation. It was no longer necessary to focus on trends in fertility, since fertility was no longer trending. At most, there was an interest in timing and spacing, as opposed to completed fertility:

"The focus of economic analyses of fertility is beginning to shift from the study of ‘completed fertility’ (the total number of children born to a woman) to that of the ‘tempo of fertility’ (the distribution of births over a woman’s lifetime). The main reason for this change of emphasis is that, at the very low levels of completed fertility now reached by most western countries, fluctuations in the ‘period’ birth rate (births per unit of time) are overwhelmingly associated with changes of tempo." \(^7\) (Cigno

\(^7\) See, for example, Hopflinger (1984), Munoz-Perez (1986).
and Ermisch, 1989).

What is the revolution which has been occurring in the economics of female labor supply? It centers on the effects of the female wage, which has been treated as a major driving force in fertility and female labor force trends since the work of Mincer (1962, 1963) — and especially since the work of Butz and Ward (1977, 1979a).

Finis Welch, in a preface to Jim Smith's study on female labor supply (1980), referred to two earlier revolutions:

"Labor economics has undergone (at least) two revolutions in the last two decades. First, there was the explicit recognition that human skills are both malleable and durable... Second, labor economists discovered women!"

What labor economists had discovered about women was that their labor supply function appeared to be very different from that of males. This finding was not easily accepted, however, so that work proceeded at a near feverish pitch throughout the 1980s to understand the underlying reasons for the observed differences. The work reported in Smith (1980) can be seen, with hindsight, to have foreshadowed later developments (particularly in the work of Heckman, Hanoch and Cogan in that volume), but it was not until the mid 1980s that the true outlines of the new revolution became apparent.

Killingsworth and Heckman (1986) provide a survey of more recent work on female labor supply in which they state:

"To motivate this discussion, we note at the outset that the results of some recent empirical studies of female labor supply differ appreciably from those of research conducted through the early 1980s. There has been a consensus of relatively long standing that compensated and uncompensated female labor supply wage elasticities are positive and larger in absolute value than those for men. In contrast, some recent studies appear to show that the compensated and uncompensated wage elasticities of women workers are little different from those of men; indeed, in this work, the female uncompensated elasticity is often estimated to be negative.

"...the results of studies of U.S. and Canadian data by Nakamura and Nakamura (1981), Nakamura, Nakamura and Cullen (1979), and Robinson and Tomes (1985). Here, the uncompensated elasticity of labor supply with respect to wages is negative (so much so that even the implied compensated elasticity is also negative in some instances). Similarly, Smith and Stelcner (1985) and Stelcner and Smith (1985) obtain uncompensated (and compensated) elasticities that, although positive, are very small in magnitude.

"Six years ago, Heckman, Killingsworth and MaCurdy (1981, p. 108) commented that elasticity estimates obtained using recently developed econometric techniques had increased the mean of what might be called the "reasonable guesstimate" of the wage-elasticity of female
labor supply. Work since then seems to have reduced the mean and substantially increased the variance of this guesstimate."

Mroz (1987) presents a more systematic explanation of why the calculated elasticities changed so dramatically:

"In this study we find that economic and statistical assumptions can have a substantial impact on the estimates of the behavioral labor supply parameters. The three most important assumptions are (i) the Tobit assumption to control for self-selection into the labor force,\(^8\) (ii) the exogeneity assumption on the wife's wage rate, and (iii) the use of the wife's labor market experience as an instrumental variable to control for the endogeneity of the wife's wage.

"The range of labor supply estimates that we fail to reject suggests that the labor supply behavior of working married women matches the estimated behavior of prime aged males. Such a conclusion conflicts with most commonly held beliefs about female labor supply. . . In this study, we are able to obtain large estimates of the income and wage coefficients. Our statistical tests, however, emphatically reject the economic and statistical assumptions needed to obtain these large wage and income effects."

Nakamura and Nakamura (1985) are also sensitive to the possible endogeneity of a woman's educational attainment, as they explain:

"We have already mentioned the possibilities that employers may use education as a screening criterion and that an education variable may act as a proxy for variables such as ability or socioeconomic background. (See Thurow, 1972, 1975; Gintis, 1971; and Fraker, 1984.) There are also those who argue convincingly that education affects tastes and hence, that education should be included as an explanatory variable in the hours of work equation for married women, in contrast to the usual practice of assuming a priori that education will only affect the hours of work of a married woman indirectly through the impact of education on the wife's wage rate. (See, for instance, Michael, 1973). The main reason for not doing this in studies like Nakamura, Nakamura and Cullen (1979) and Nakamura and Nakamura (1981, 1983a, 1983b) is that the predicted values of the wage rate that enter the hours equation are strongly affected by education, leading to severe problems of multicollinearity if education is also introduced into the hours equations.

\(^8\) Killingsworth and Heckman (1986) explain this statement as follows: "The Tobit specification . . . implicitly assumes that hours of work vary continuously from zero (at a wage equal to the reservation level) to progressively larger positive amounts (at wages greater than the reservation level), with no jumps or discontinuities. In contrast, the generalized Tobit specification (sometimes called 'Heckit') implicitly allows for a discontinuity in labor supply at the reservation wage such that hours worked are zero below the reservation level and some large amount above the reservation level. The latter approach has sometimes been characterized as a means of allowing for the labor supply discontinuities that may be induced by fixed costs of labor market entry. [For further discussion see Cogan (1980b) and Killingsworth (1983, esp. pp 141-148).]"
as an independent explanatory variable. We were not even able to get around this problem by using a nonconventional instrument for the wage variable in the hours of work equations (Nakamura and Nakamura, 1983). This problem may be circumvented to some extent, however, in studies like this one that control for previous work behavior and that therefore control to some extent for education-related circumstances and tastes for work that are reflected in the person's previous work behavior."

More recent studies, such as that of Smith and Stelcner (1988), continue to produce only small wage and income effects on female labor supply, at best.

F. What Does This Mean for Fertility Research?

These developments in female labor supply underline the naivete with which we have been approaching neoclassical models of fertility. They indicate the need to examine our assumptions of exogeneity of the female wage in fertility equations -- especially in equations which do not contain any controls for education or work experience. Just as the response of female labor supply to female wages was overestimated, the response of fertility to female wages may also have been overestimated -- or mis-estimated. But there has been less motivation for researchers to pursue this line of enquiry: fertility findings using this simplistic model conform to priors based on the neoclassical model, whereas researchers working on the economics of female labor supply were to some extent driven by what they saw as anomalies between male and female labor supply functions.

"Of course, there is no such thing as a distinct 'model of female labor supply' per se: any theory worthy of the name ought to be just as applicable to men's as to women's labor supply" (Killingsworth and Heckman, 1986, p. 126).

This paper does not argue for a radical departure from the neoclassical model. Rather, it argues for a suspension of the priors with which that model is approached: in particular, the prior assumption that the price effect of the female wage will strongly dominate its income effect. Simply removing -- or even temporarily suspending -- that prior assumption will enable researchers to approach the fertility equation with greater motivation to question the exogeneity assumptions inherent in the standard, simplistic neoclassical model. At the aggregate level, findings of 'causality' from the female wage to fertility may simply have reflected the fact that decisions taken early in a woman's life -- to postpone/reduce childbearing and pursue a career, and hence to invest in human capital -- resulted in both increased wages and reduced fertility. This approach is adopted in the work of Blackburn, Bloom and Neumark (1990), where it is demonstrated that, conditional upon knowledge of the timing of first birth, late childbearers tend to invest more in human capital, thus exhibiting higher wages than early childbearers. Similarly, Lehrer and Nerlove (1986) prepared an excellent survey of the literature on fertility and female labor force participation, in which they refer to the work of Lehrer and Stokes (1985) in emphasizing that
"Those women who at a young age foresee low fertility and a strong attachment to the labor force are more likely to make investments in human capital specific to the labor market than are their counterparts with weaker commitment to market activities." (p. 188).

This is consistent with the Easterlin model, in which reduced male earning potential relative to aspirations caused couples to postpone/reduce childbearing and household formation, and women to pursue careers to supplement family income. The longer term aggregate effects of these decisions would be increased female earnings (uncorrected for education and experience) and reduced fertility. Results in Macunovich (1990) using both Butz and Ward's time series data, and revised versions of that data from the Current Population Survey, indicate the strong likelihood that this was in fact the underlying model. Lags of the dependent variable \(^9\) led to a much better 'fit' of the data, indicating that in fact the true direction of causation was not from the female wage to fertility. At best, there were feedback mechanisms operating. Similarly, Michael (1985, p. S119) indicates that 'the findings support the hypotheses that lagged values of fertility (and, marginally, divorce) are correlated with married women's LFPR.' These findings can be interpreted as indicating (as Easterlin maintains) that changes in both fertility and FLFP were caused by changes in relative income, with a longer lag period for FLFP than for fertility.

As a result, it can be seen that neglecting to control for female education and female labor force participation (FLFP) in a fertility equation can bias the coefficients on both the female wage and on the measure of husband's income — whether relative or absolute. This problem was demonstrated in Macunovich (1989), where it was shown that the inclusion by Butz and Ward of interactions of FLFP with the female wage and with male income, while excluding the FLFP measure on its own, was the probable cause of the 'perverse' sign on their variable \(K \cdot \ln Y_m\). \(K \cdot \ln Y_m\) was an interaction term used by Butz and Ward which was expected to indicate the differential between the effect of male income in households with a working wife (where a pure income effect was hypothesized) and the effect in households with a non-working wife (where the effect was a combined income and substitution effect because of the male income's function as a proxy for the shadow value of her time). However, because their variable \(Y_m\) was absolute rather than relative income, all other things being equal, the households with working wives must have had higher (unobserved) values of parental income than the households with non-working wives, if the relative income hypothesis holds. Including the FLFP variable on its own would have proxied that omitted variable — its exclusion biased the effect of \(K \cdot \ln Y_m\).

---

\(^9\) Three year lags are shown in the paper, and with the CPS data it was found (in results not presented in the paper) that the fit improved successively with lags up to ten years.
1. The Need to Control for Education in a Fertility Equation

In a similar fashion, to the extent that the female wage is a function of education and experience, and to the extent that education and experience (labor force participation) have direct negative effects on fertility, including the female wage in a fertility equation without the other two variables will produce a downward bias on the coefficient of the female wage. The expected direct negative effect of education on fertility has been argued by many in the literature. It is effectively a correlate with the desire for increased child quality argued by Becker (1981), and Rosenzweig and Schultz (1989) demonstrate with U.S. data that couples with higher levels of education have a wider knowledge of contraceptive methods, and are more efficient especially in using methods ‘for which there is little information and large scope for misuse’. Thus, even with a given level of preference for children, individuals with higher levels of education will tend to have lower levels of fertility. In addition, Lehrer and Nerlove (1986) argue persuasively that unobserved taste effects operate through the education variable:

"Nerlove (1974) has noted that the coefficient on female education in a fertility equation may overstate the impact associated with the opportunity cost of time. Positive assortative mating by education in the marriage market leads men with higher-than-average levels of education to marry women who also have higher-than-average levels of schooling. Differences in taste are not likely to be reflected in the educational attainment of men; however, it is very plausible that women with limited interest in market activities and high preferences for children will, in general, seek less formal education than women with opposite preferences. If, as is likely, there is positive assortative mating by preferences for children, then men with a given educational attainment with strong preferences for children will tend to marry women with less schooling than the average associated with the level these men have achieved. If the husband’s schooling level is associated primarily with an income effect, while his wife’s education is associated mostly with the cost of time, it follows that the negative impact of the opportunity cost of time as measured this way on fertility will tend to be exaggerated."

Thus, the results of these studies indicate that it would be inappropriate either to use education on its own as a proxy for the value of a wife’s time in childrearing, or to include the wage on its own (to the extent that it is a function of education) without controlling separately for the negative taste effects of education. More work is required on these direct effects of education, controlling separately for labor force participation and wage, especially in light of the results of Danziger and Neuman (1989), who find (consistent with Ben-Porath, 1973) that the effects of education tend to be U-shaped, with the minimum occurring at about 16.5 years of education. It is unclear, however, since they estimate only linear wage effects and do not control for costs of childcare, whether this increasing effect of female education levels over 16.5 is an effect of education per se or an increasing income effect of increasing female wages.

2. Effects of Purchased Childcare
This last statement leads into a perhaps even larger question regarding the ability of women to purchase childcare services, and subsequent implications of that fact regarding income effects of the female wage on fertility. This is a factor which has been surprisingly neglected in the literature to date\(^{10}\), although recent work by Ermisch (1989) has produced a model of family size decisions incorporating substitutability between mother’s time and purchased services for childcare. Using this model, he demonstrates quite effectively that

"The effect of a higher price for child care, \( p \), on family size is unambiguously negative. . . There is a tendency for the impact of mother's wages on family size to vary with the level of wages and the price of market child care. . . While by no means certain, there is, therefore. . . a tendency for the impact of mother’s wages on family size to become less negative (or more positive) at higher levels of wages."

(pp. 83-84)

Ermisch goes on to demonstrate that this is in fact the case in a cross section of data for Britain in 1980. He presents evidence that the average wage of women who purchase childcare is significantly higher than the average wage of women who do not purchase such care, and that the effect of the female wage is positive for women in his sample ‘having a potential earning of £2.20 per hour or more at marriage (about 5% of women in the sample).’ (p.95) He also demonstrates a backward bending labor supply curve for women with higher wages.

It is reasonable to expect that this effect would be even greater in the U.S., where studies have shown a high propensity toward the use of purchased childcare. For example, Lehrer (1989) points out that the 1982 National Survey of Family Growth indicates that 59.5% of all preschoolers of working mothers receive paid childcare — and in the 40.5% of cases in which preschoolers are taken care of by relatives, 39% involve payments of some type. By way of comparison, Joshi (1990) states that "mothers’ foregone earnings remains particularly relevant in Britain since there is very little recourse to paid child care and mothers take prime responsibility for unpaid child care." (p.42)

O’Connell and Bloom (1987) look at the lack of low-priced childcare as an impediment to labor force entry of mothers with young children and find that

"About one-fourth of women not in the labor force (or 13.4 percent of all women) who had children under 5 years old said they would look for work if child care were available at a reasonable cost. This figure contributes to a ‘potential’ labor force participation rate of about 62 percent for women with preschool children in 1982 (i.e., the sum of 48.1 percent of women with preschoolers who were actually in the labor force and 13.4 percent with preschoolers who said they would look for work). The potential labor force participation rates for maritally-disrupted women and for single women were 76 and 74 percent respectively,

\(^{10}\) Although the issue of childcare has been raised, it has not until recently been brought directly into analyses of fertility. Rather, as in Stolzenberg and Waite (1984) and Cramer (1979) it has been examined in terms of its effect on maternal employment.
Here we see direct evidence that a higher wage would permit more women to combine working and childrearing: an income effect of the female wage on fertility. And Cramer (1979) demonstrates that the opportunity cost (measured as number of hours lost times the wife's wage) of babies in the U.S. in 1967 was not dependent on the wife's potential wage, indicating that while lower-wage women lose substantial hours to childcare, higher wage women lose fewer because of the use of purchased childcare.

However Ermisch, like Schultz (1976), assumes that in time series data we would not expect to see these effects of changing costs of child care relative to wages, since 'average women's wages and the price of child care move together over time.' (p.85) Lee (1986) has questioned this assumption, given the greater potential for economies of scale in childcare than in women's jobs generally, and Blau (1989) demonstrates a relatively elastic supply of labor in the childcare market which has prevented subsidies from driving up the cost of childcare.

3. The Need to Control for Work Experience in a Fertility Equation

As with education, female labor force participation itself might exert direct negative effects on fertility, separate from the effect of the female wage. This occurs for at least two reasons: unobserved taste effects which might influence both a woman's preferences for labor market activity and her preferences for children, and 'state dependence' effects such as those identified by Heckman (1979) and Eckstein and Wolpin (1989), among others. These state dependence effects might be explained as an increased propensity of a woman to continue participating in the labor market because of the effects of current and past levels of participation on her wage, and consequently on the fixed costs of labor market entry if she were to leave the labor force. The more she participates, the higher is her offered wage, and the greater would be her human capital depreciation if she were to leave the labor market (and hence the greater would be her expected wage loss if she planned on returning).

Because her wage reflects these effects of her current and past labor force participation, it is necessary to include some measure of a woman's labor force participation along with educational attainment and the wage in a fertility equation in order to control for these non-wage effects. This has been done in various studies of female labor supply, where the effect has been drastically to reduce the estimated effect of the female wage on her labor supply, as discussed in earlier sections of this paper.

If the wage effect on female labor supply is very small, insignificant or even negative, as indicated by recent studies, and if unpaid care with relatives is prevalent among lower wage females and paid care is dominant among higher wage females, as indicated by Lehrer (1989), then it stands to reason that the female wage should not exert much of a negative impact at all on fertility, if proper controls are maintained for 'taste' effects. However, the degree of a woman's labor force participation (in terms of hours and weeks worked) might be expected to exert a negative influence, to the extent that higher levels of participation are incompatible with available methods of purchased/arranged substitute childcare.
This is the approach taken in work by Macunovich and Lillard, as reported in Macunovich (1989), a study which uses PSID data for the period 1968-1985 to examine marital fertility in the first birth interval for 2,078 marriages in a hazard model framework. Here, an attempt is made to test the Butz and Ward model at the micro level\textsuperscript{11}, and then to examine the effects of adding controls for the extent of the wife’s labor force participation. Although no model of ‘pure’ Butz-Ward effects is presented in the paper, it can be seen there that when controls for the wife’s educational attainment are included, the effect of the female wage is positive but insignificant, and when controls for the degree of her labor force participation are included as well, the effect of the wife’s wage becomes significantly positive. The effect of her degree of labor force participation, however, is negative. This latter effect is strongly significant for women who work 48 weeks or more per year, and weakly (at the 10\% level) for women who work 1-16 weeks per year. For women in the middle range (17-47 weeks per year), the effect is not significant.

These results can be interpreted as suggesting that measures of the degree of female labor force participation can pick up ‘taste’ effects, combined with a nonlinearity in the effect of the female wage. That is, assuming that the category of women who work 48+ weeks per year can be expected to contain a higher than average proportion of ‘career-oriented’ women, the coefficient on this category will be biased downward by ‘taste’ effects. In a similar fashion, the observed negative effect of part-time work on female wages might cause the coefficient on this category to be downward-biased because of larger substitution effects among low-wage women.

These results of Macunovich and Lillard are certainly not definitive, but they do indicate the scope for further work in identifying the ‘true’ effect of the female wage on fertility. In addition, they indicate the importance of determining this true effect, for fertility forecasting. For example, simulations presented in the paper demonstrate that the income effect of the female wage is so strong, that the likelihood of conception will actually increase for a woman as a result of entering the labor force at the median wage, if she works less than 48 weeks per year (by 18.9\% if she works 1-16 weeks and by 37.2\% if she works 17-47 weeks).

Work by a number of other researchers provides further justification for the use in fertility equations of variables measuring the intensity of female labor supply. For example, the criticisms outlined in section F of this paper regarding the Tobit assumption of a continuous distribution of hours

\textsuperscript{11} One model in the paper is described as an ‘unaugmented’ Butz-Ward model. In this model, the coefficients on male income and the (working) wife’s wage are positive and insignificant, and the differential effect of the male income in one-earner families is positive (indicating a stronger income effect in one-earner than in two-earner families), contrary to expectations. However, as pointed out by participants in Columbia’s Labor Economics Workshop, this ‘unaugmented’ model in fact contains controls for female education (unlike the Butz-Ward model), which exert a significant negative effect on the probability of a birth. In Butz and Ward’s interpretation, this negative effect of female educational attainment would be an indication of the negative effect of the wife’s opportunity costs. However, as argued in this paper, interpreting the coefficient on the wife’s education as a measure of the price of time effect severely overstates the latter, to the extent that education reflects unobserved ‘taste’ variables.
of work indicate the possible relationship between hours worked and choice/availability of child care arrangements. Researchers have found that it is necessary to recognize the possibility of a discontinuity — the fact that because of fixed costs of labor market entry (i.e., child care costs) a woman enters with some large, positive number of hours rather than with zero hours as is assumed by the Tobit specification (Hanoch, 1980; Cogan, 1980; Mroz, 1987). As a result, as Mroz indicates, 'the hours of work decisions made when the woman is in the labor force appear quite distinct from her labor force participation decision' (1987, p.790). Blank (1988) demonstrates that

"An estimation system that separates the labor force participation decision from the weeks and hours of work decisions is shown to perform well. With labor-force participation accounted for separately, the determinants of weeks and hours are quite different from those estimated by independent or bivariate Tobits. . . Race and low education negatively influence the probability that an individual will work, but once a work decision is made, these variables now have positive effects on both the number of hours and the weeks of work that occur. . . Most surprising, having children under the age of six makes a woman less likely to work and likely to work fewer weeks but increases the hours she works. . . A large number of dependents increases the probability of work but has negative effects on weeks and hours."

These findings demonstrate that the issue of substitutability between a woman’s time in the home, and purchased goods and services, makes it necessary to consider the degree of a her labor force participation separately from the actual decision to participate. This is particularly important as we begin to work with datasets in which the vast majority of women all have some labor force experience in a given year. The (1/0) labor force participation variable imparts no new information of its own, whether it is a function of exogenous wage changes of the husband’s (relative) income, or both. Any 'new' information is probably caused simultaneously with fertility and education by 'unobservables' (ability, motivation, career-orientation), so that a statistical relationship would be spurious.

However, the degree of a woman’s participation — in terms of hours and weeks worked — seems to provide a new source of independent information. Blank’s (1988) evidence lends some credence to this line of thought, given that different variables are involved in the two sets of decisions, and that the same variable might differ in the sign of its effect in the two decisions. But even more, she indicates evidence of ‘labor-market rationing’: ‘unemployment and underemployment explain a significant proportion of the remaining negative residuals in the equations for weeks and hours of work.’ This result is corroborated by Sundt (1988):

"The probability that a woman desires full-time work but is unable to find it is as high as 16.5% for young women, and 11.5% for minority women."

In addition, Ilmakunnas and Pudney (1990) find that in Finland, "one in five women is seriously
constrained in her choice of hours of work." They go on to demonstrate that

"Barriers to choice arise mainly from a shortage of part-time opportunities, particularly at higher wage levels. This tends to force women out of the labour force or into full-time work, from their desired option of part-time employment. This generates a spurious correlation between observed hours and wage rates, which may produce an upward bias in conventional estimates of wage elasticity."

Thus there may be a large exogenous component in the degree of a woman’s labor force participation, which should be controlled for in modelling fertility — and especially in estimating the effect of the female wage on fertility.

G. Do These Models Adequately Reflect Gender Roles in Our Society?

Deconstructivists have shown us that it is perhaps as important to consider the context in which a model is formulated, as it is to consider the model itself. Both the New Home Economics and the Easterlin models have been attacked as chauvinistic: the New Home Economics model for its reliance on arguments of comparative advantage in explaining the sexual division of labor, its assumption that wives enter the labor force only to earn extra ‘pocket money’ and its assumption of a household utility function which jointly maximizes the utility of all household members (Ferber and Birnbaum, 1977 and 1982; Manser and M. Brown, 1979; L. Brown 1984); and the Easterlin model for the fact that

"The basic implication seems to be that women are programmed primarily as wives and mothers and that satisfaction of these roles is periodically frustrated by economic conditions, which in turn are influenced by cohort size. The argument does not allow for changes in the status, attitudes, and interests of women who may be refocusing their lives around the same occupational and nonfamilial interests traditionally associated with the male world." (Westoff, 1986, p.166)

Although knowledge of such allegations provides important insights into the acceptance — or otherwise — of a model, it is important to differentiate between models which are sexist, and models which describe a sexist society. The latter type of model, so long as it takes account of any trends in sex role norms and attitudes, can still be expected to perform well as a forecasting tool, even if it is unacceptable as a normative model.

The work of Matthaei (1982) and C. Brown (1985, 1987) have added considerable insights regarding the important contribution made to household income by the working wife’s wage, and Manser and M. Brown (1979) have set out a framework for analyzing bargaining within the household context, especially in terms of the relative power of the respective bargainers. Although work on the latter topic is only in its infancy, findings presented in previous sections with regard to the former indicate a serious need in the New Home Economics model to take more account of possible income effects of the female wage.

With regard to sex role attitudes as they affect the relevance of the Easterlin model, as Crimmins,
Easterlin and Saito (1990) demonstrate, the data do not indicate substantial changes in the underlying preferences of young adults toward children and family formation. Their findings at the national level are supported by findings from surveys of undergraduates at Williams College in 1990 indicate a preponderance of young women who, although they intend to pursue a career, also intend significantly to reduce their labor force participation during childrearing years. They expect on average to work 44 hours per week in the labor force prior to bearing children, but this drops to only 17.67 when children under three are present. Not surprisingly, perhaps, males in the sample expect their wives to participate even less in the labor market during the childbearing and rearing years. Only 4% of the sample expect not to marry, and the average expected family size is 2.5 (2.4 for females, and 2.58 for males). 87 percent of the women, and 95 percent of the men surveyed expect to have two or more children. (Goodwin, 1990; Kwon, 1990; Laferriere, 1990). These findings underscore Easterlin’s contention that underlying traditional sex role attitudes have not changed significantly over the past few decades.


<table>
<thead>
<tr>
<th>Year</th>
<th>1968</th>
<th>1980</th>
<th>1989</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professional, excl. nurses and non-college teachers</td>
<td>3.94%</td>
<td>5.01%</td>
<td>6.24%</td>
</tr>
<tr>
<td>Nurses, librarians and non-college teachers</td>
<td>16.33</td>
<td>16.77</td>
<td>15.43</td>
</tr>
<tr>
<td>Proprietors, managers and administrators</td>
<td>3.24</td>
<td>8.33</td>
<td>11.35</td>
</tr>
<tr>
<td>Clerical</td>
<td>35.16</td>
<td>33.44</td>
<td>28.35</td>
</tr>
<tr>
<td>Sales Workers</td>
<td>7.00</td>
<td>8.32</td>
<td>12.15</td>
</tr>
<tr>
<td>Craft and kindred including technicians</td>
<td>1.29</td>
<td>2.29</td>
<td>3.36</td>
</tr>
<tr>
<td>Operatives</td>
<td>15.94</td>
<td>10.83</td>
<td>8.26</td>
</tr>
<tr>
<td>Laborers</td>
<td>1.84</td>
<td>2.36</td>
<td>1.74</td>
</tr>
<tr>
<td>Service:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household</td>
<td>3.71</td>
<td>1.28</td>
<td>1.10</td>
</tr>
<tr>
<td>Non-household</td>
<td>11.56</td>
<td>11.38</td>
<td>12.02</td>
</tr>
</tbody>
</table>

Finally, examination of the actual distribution of women by occupation, as set out in Table 1, leads one to believe that our preconceptions are too often influenced by our own experiences, and those of our acquaintances. While the share of women in this age group in the operative group declined by nearly 8% between 1968 and 1989, and it is encouraging that approximately 10% more women aged 20-39 were in the professional and managerial groups in 1989, than in 1968, the decline in the proportion in the clerical group has been balanced by increases in sales and service. In addition, Maxwell (1990) reports that

12 Table 1 is drawn from tabulations of the March Current Population Survey Public Use Tapes.
"College-educated males exhibit the greatest occupational mobility with their movement into management positions. While females, irrespective of educational level, are overwhelmingly employed as clericals... the movement of high-school educated females into the clerical field suggests that these females will spend nearly half their work lives as clericals."

It should also be pointed out that Easterlin's relative income model accords very well with hypotheses put forward by a number of feminist writers regarding the cause of observed increases in female labor force participation in the twentieth century. (See, for example, Clair Brown, 1985, 1987 and Matthaei, 1982). Their hypothesis is that

"The development of an expanding system of commodity needs and the creation of constant neediness had a powerful effect on homemakers. Not only was the homemaker transformed from a home producer to a consumer of commodities, she was the family member most caught up in competitive consumption, for it was her job to see that the family's needs were fulfilled...

"Faced with this family need for income, some homemakers interpreted the dictates of vocational homemaking in a new and heretical manner. Rather than adjusting their families' needs to their husbands' incomes, or patiently waiting for their families' incomes to increase, these homemakers judged their husbands' incomes to be inadequate to their families' commodity needs and entered the labor force to remedy the situation." (Matthaei, 1982, pp. 245-248).

H. How Well Has the New Home Economics Model Predicted Fertility Trends?

This is not as simple a question as it might seem. On the basis of the time series produced by Butz and Ward (1977, 1979a), fertility was forecast to continue falling sharply throughout the 1970s and into the 1980s. As indicated in section D, by 1983 those forecasts were underpredicting observed fertility by over 30% for women aged 25-34, and by nearly 25% for women aged 20-24. However, as we saw in section D, the time series estimated by Butz and Ward for the female wage was inaccurate, and replacing that estimated series with one derived from CPS microdata produces an insignificant coefficient on the female wage in the Butz-Ward model. But Smith and Ward (1984, 1985) question the validity of average female wages recorded in the CPS, as an accurate measure of the opportunity cost of women's time in childrearing. They, together with Goldin (1989) and Heckman (1979) present evidence of heterogeneity in the female population which implies that secular increases in female labor force participation have been extensive rather than intensive, so that increases at the margin tended to reduce the aggregate average level of experience in the female labor force, and thus recorded aggregate female wages.

As indicated in section D, Smith and Ward (1984, 1985) found significant negative effects of their estimated female wage on fertility, and they predict further strong increases in female wages through the
turn of the century:

"Women's wages in the workforce over the next twenty years will be far different from the past. We project that the skills of the typical female worker will increase sharply relative to males by the year 2000. First, in contrast to past trends, women's education will increase faster than men's. In addition, market participation is now rising more rapidly among more educated women than among less educated women.

"The average work experience of both the female workforce and the population will also increase significantly. By the year 2000, a 40-year-old working woman will have 5.2 more years of work experience than her counterpart had in 1980. As a result, we estimate that wages of working women will rise at least 15 percent faster than those of men over the next twenty years." (Smith and Ward, 1984, p.xiv)

On this basis, it seems safe to say that proponents of the New Home Economics model still predict declining fertility, which is at odds with the observed pattern since 1974. It is not possible to substantiate this statement, however, since Smith and Ward have not produced forecasts of fertility using their estimated equations — nor have they provided their estimated time series of wages in the female population, or information on deviations of estimated from actual fertility in the forecast period.

This leaves us with an anomaly in the New Home Economics model, then. Consistent with the model's hypotheses, Smith and Ward find a strong negative effect of the female wage on fertility, and a strong positive effect of the female wage on female labor supply. During the period covered by their data, they demonstrate increases in the female wage and we observe continuing increases in the level of female labor force participation, consistent with the hypothesis, but a total fertility rate which remained relatively unchanged from 1974 into the 1980s, and then increased sharply. This would seem to indicate either that their wage series is inaccurate (but then we're left without theoretical underpinnings for the increased female labor supply) — or that there are shortcomings in the model.

Based on the findings set out in previous sections, it seems most likely that the effects of the female wage have been mis-estimated in Smith and Ward's work. It will be noted in the quote from Smith and Ward above, that they use female wages unadjusted for education and experience, so that their findings of a negative effect of female wages on fertility must be interpreted as a correlation rather than demonstration of a causal relationship, which is strongly influenced by negative taste effects operating through education and labor force participation, and hence through the female wage.

Given the findings of Macunovich and Lillard as reported in Macunovich (1989), there is an explanation for the divergence between the fertility rate and the female participation rate in the 1970s and 1980s, which is consistent with the New Home Economics model so long as one is willing to suspend the prior assumption of a dominant price effect of the female wage. If the negative effect observed by Smith and Ward is in fact working through education and career choice, then it is possible that as the female wage has risen, it has begun to exert an increasingly strong income effect, given women's ability to purchase childcare services. This positive income effect of the female wage, in combination with a
negative effect of education and labor force participation, would have produced the essentially dormant fertility rates we observed from 1974-1985, and the increases we have observed since then.

I. Why the Baby Bust Cohorts Haven't Boomed Yet

A critic might well point to the simple fact that fertility rates remained dormant through most of the 1980s, despite the entry of the peak baby boom cohorts into the labor market in the late 1970s — as adequate cause to reject the Easterlin relative income hypothesis. However, four possible factors might be suggested now, with hindsight, to have been responsible for mistaken expectations of an earlier upturn. They are:

1) insufficient attention to the effects of immigration, especially illegal immigration -- on the feedback between relative cohort size and relative income.

2) failure to give adequate consideration to the contribution of female earnings to total household income in baby boomers’ households, and the possible need to include this source of income in relative measures for subsequent baby bust households;

3) the neglect of reinforcing ‘bandwagon’ effects of changes in fertility and female labor force participation, on social norms and attitudes;

4) The assumption of symmetry in the direct labor market effects of cohort size pre- and post- peak;

Point (1) was touched on earlier in this paper, and is also discussed in Wachter (1980) (and in all of Easterlin’s work, although it seems to have been overlooked there by most readers): it is sufficient simply to reiterate that smaller cohort sizes will not be translated into favorable earnings differentials if there are other supply-side effects occurring. In this case we have a ‘leaky’ system, rather than a closed one, so that the beneficial effects of smaller cohort sizes must be expected to manifest themselves with a lag, if at all.

Point (2) was introduced by Oppenheimer (1976) in her revised version of the Easterlin model, but the asymmetric effects do not seem to have been appreciated generally in the literature. That is, the use of husband’s earnings in pre-baby boomer households might have been an adequate proxy for total household income (and therefore for material aspirations of the boomers growing up in those households). However, the high levels of female labor force participation, and importance of female earnings as a component of total household income in baby boomer households, make it imperative that baby bust cohort earnings be viewed relative to parents’ total income. As a result, this is another reason why it is to be expected that relative income will be slower to rise for persons on the trailing edge of the boom, just as it was slow to fall for persons on the leading edge.

Point (3) has to do with the norms and attitudes discussed by Westoff (1986), and others. It has been argued that the feminist movement received its primary impetus from the rapid increase in female labor force participation (Mason et al, 1976), and that the increasing acceptance of working mothers has occurred as a result of ‘cognitive dissonance’: individuals adjusting their beliefs to accord with their
experiences. In this case, then, an initial increase in female labor force participation, and reduction in fertility, which were motivated by relative income pressures, created reinforcing effects which in Wachter's (1988) terminology resulted in a ‘cascade’ model, with labor force participation increasing and fertility decreasing at rates higher than would have been predicted with a simple relative income model.\footnote{Wachter acknowledges discussions with Paul David and Andrew Cherlin, in motivating his work with this concept.}

Wachter (1988), working with highly abstracted mathematical simulations, has experimented with the dynamics of cycles when such bandwagon or cascade effects are explicitly considered -- the increased momentum generated once a turning point has finally been reached, and the inertia which builds up in approaching that turning point. He finds that when such effects are included, ‘this model can sustain cycles of proper period.’ If this is the case, then norms and attitudes might take a pendulum swing back in the other direction, but almost certainly there will be a lag effect for early baby bust cohorts -- especially if point (2) has any validity.

Point (4) is the subject of study in Macunovich and Lillard (1989), which demonstrates that attempts to quantify asymmetric cohort size effects can be extremely enlightening. The authors build on an established tradition of research which has found strong negative effects of cohort size on white male earnings.\footnote{See, for example Welch (1979), Ahlburg (1982), Anderson (1982), Berger (1984, 1985, 1989), Tan and Ward (1985) and Dooley and Gottschalk (1985)} However, the research by Macunovich and Lillard was conducted using data for all males (controlling for race) and more importantly it included a ‘cohort size difference variable’ which, since it measured the five year change in cohort size, was positive for cohorts on the ‘leading edge’ of the baby boom, and negative for cohorts on the ‘trailing edge’. This variable was found to be highly significant (along with the original cohort size variable) in its effects not only on male wages, but also on employment probabilities, unemployment, and hours and weeks worked.\footnote{Berger (1989) also looks at effects of position in the demographic cycle, but does so without differentiating between cohorts on the leading and lagging edge, so that it is somewhat difficult to interpret the signs on his coefficients. Also, because his three cohort variables are identical except for lags, there is the possibility of collinearity among them. Nevertheless, he also finds position in the demographic cycle an important covariate.}

The sign of the cohort difference variable in these equations indicated that there are ameliorating effects on cohorts in the leading edge and exacerbating effects on cohorts in the trailing edge: effectively, we observe the equivalent of a bottleneck effect as the peak of the baby boom enters the labor market. Like the pig in the python, this large influx of new entrants must be digested by the market before it can accommodate the cohorts of the trailing edge. In particular, there is a lag before the beneficial effects of smaller cohort size are registered in labor market outcomes for these smaller cohorts.

Simulations incorporating these regression outcomes indicated that the effective lag was approximately five years: wages which we would otherwise have expected to begin rising in the early
1980s, as a result of declining cohort size per se, did not begin rising until about 1985, all other factors held constant at their 1985 levels. These simulations fit the observed pattern of entry level male wages extremely well, over the simulated period. In addition, the simulations indicate a continued rise in young male wages through the end of the century.

These findings are extremely important for the relative income model, since they document one reason for the delay in the predicted increase in fertility in the U.S. Figure 3 indicates the actual path of relative income in the U.S. since the early 1970s, which turned around at about 1985. It presents the ratio of a five-year moving average of the real earnings of males with 1-5 years of labor market experience, to a five year moving average of the real earnings of males with 25-29 years of labor market experience which has been lagged five years. The results are not noticeably sensitive to the choice of experience group in the denominator: the use of various combinations of five year groups from 20-24 through 35-39 produces virtually identical results. The source of the data for this analysis is the Public Use Tapes of the March Current Population Survey, 1968-1989.
J. An Interpretation of Recent Fertility Trends

In light of all of the foregoing, it seems that a strong case can be made for an interpretation of recent fertility trends which incorporates both relative income and female wage effects. As indicated in the previous section, relative income in the U.S. began a strong upward trend around 1985, which can be explained in terms of relative cohort size effects once we accept the asymmetry of such effects as argued by Macunovich and Lillard (1989). This can be expected to exert a positive effect on fertility, and a negative effect on female labor force participation. Tabulations from the March Current Population Survey tapes support the latter hypothesis, in that levels of female participation within age- and education- specific groups appear to have 'topped out' over the last three years, and even to be declining in some groups, as indicated in Table 2. And, as was indicated in Figure 1, the total fertility rate appears to have exhibited a strong upward trend in the last three years.

In a similar fashion, the lack of any apparent trend in the total fertility rate in the period from 1974-1985 can be explained in a relative income framework. That is, although the combination of falling relative (and even absolute) incomes of young males and rising female labor force participation might have been expected to produce a continuing decline in fertility rates during this period, an increasingly dominant income effect of female wages (on the basis of Macunovich and Lillard’s findings for the period 1968-1985, as reported in Macunovich, 1989) could have mitigated such a downward tendency in the fertility rate.

For the future, in the absence of further significant increases in the female participation rate, we


<table>
<thead>
<tr>
<th></th>
<th>&lt; 12 yrs educ</th>
<th>12-15 yrs educ</th>
<th>16+ yrs educ</th>
<th>all</th>
</tr>
</thead>
<tbody>
<tr>
<td>all women:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>0.388</td>
<td>0.658</td>
<td>0.853</td>
<td>0.635</td>
</tr>
<tr>
<td>77</td>
<td>0.451</td>
<td>0.667</td>
<td>0.851</td>
<td>0.651</td>
</tr>
<tr>
<td>78</td>
<td>0.448</td>
<td>0.681</td>
<td>0.849</td>
<td>0.659</td>
</tr>
<tr>
<td>79</td>
<td>0.460</td>
<td>0.712</td>
<td>0.866</td>
<td>0.687</td>
</tr>
<tr>
<td>80</td>
<td>0.448</td>
<td>0.696</td>
<td>0.837</td>
<td>0.670</td>
</tr>
<tr>
<td>81</td>
<td>0.433</td>
<td>0.712</td>
<td>0.843</td>
<td>0.682</td>
</tr>
<tr>
<td>82</td>
<td>0.440</td>
<td>0.711</td>
<td>0.845</td>
<td>0.686</td>
</tr>
<tr>
<td>83</td>
<td>0.412</td>
<td>0.720</td>
<td>0.859</td>
<td>0.688</td>
</tr>
<tr>
<td>84</td>
<td>0.444</td>
<td>0.719</td>
<td>0.839</td>
<td>0.690</td>
</tr>
<tr>
<td>85</td>
<td>0.490</td>
<td>0.725</td>
<td>0.873</td>
<td>0.704</td>
</tr>
<tr>
<td>86</td>
<td>0.478</td>
<td>0.727</td>
<td>0.882</td>
<td>0.706</td>
</tr>
<tr>
<td>87</td>
<td>0.463</td>
<td>0.741</td>
<td>0.882</td>
<td>0.717</td>
</tr>
<tr>
<td>88</td>
<td>0.454</td>
<td>0.737</td>
<td>0.886</td>
<td>0.714</td>
</tr>
<tr>
<td>89</td>
<td>0.445</td>
<td>0.729</td>
<td>0.880</td>
<td>0.707</td>
</tr>
<tr>
<td>married women:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>0.331</td>
<td>0.583</td>
<td>0.820</td>
<td>0.553</td>
</tr>
<tr>
<td>77</td>
<td>0.427</td>
<td>0.615</td>
<td>0.853</td>
<td>0.598</td>
</tr>
<tr>
<td>78</td>
<td>0.392</td>
<td>0.614</td>
<td>0.837</td>
<td>0.587</td>
</tr>
<tr>
<td>79</td>
<td>0.406</td>
<td>0.641</td>
<td>0.863</td>
<td>0.612</td>
</tr>
<tr>
<td>80</td>
<td>0.435</td>
<td>0.632</td>
<td>0.803</td>
<td>0.606</td>
</tr>
<tr>
<td>81</td>
<td>0.406</td>
<td>0.652</td>
<td>0.828</td>
<td>0.616</td>
</tr>
<tr>
<td>82</td>
<td>0.413</td>
<td>0.652</td>
<td>0.817</td>
<td>0.621</td>
</tr>
<tr>
<td>83</td>
<td>0.390</td>
<td>0.674</td>
<td>0.814</td>
<td>0.631</td>
</tr>
<tr>
<td>84</td>
<td>0.383</td>
<td>0.681</td>
<td>0.830</td>
<td>0.635</td>
</tr>
<tr>
<td>85</td>
<td>0.486</td>
<td>0.672</td>
<td>0.887</td>
<td>0.650</td>
</tr>
<tr>
<td>86</td>
<td>0.421</td>
<td>0.658</td>
<td>0.885</td>
<td>0.631</td>
</tr>
<tr>
<td>87</td>
<td>0.446</td>
<td>0.702</td>
<td>0.889</td>
<td>0.674</td>
</tr>
<tr>
<td>88</td>
<td>0.428</td>
<td>0.692</td>
<td>0.864</td>
<td>0.659</td>
</tr>
<tr>
<td>89</td>
<td>0.390</td>
<td>0.677</td>
<td>0.881</td>
<td>0.643</td>
</tr>
</tbody>
</table>
can expect on the basis of Macunovich and Lillard’s work reported in Macunovich (1989), that positive female wage effects would become more dominant in fertility trends -- especially if wages experience increases as a result of the effects of a labor supply 'squeeze' caused by declining cohort size in the 90s. Given the small or even negative wage elasticities of female labor supply demonstrated in recent studies reported in this paper, such wage increases would have little effect on female labor supply, but a large (positive) effect on fertility. Thus, abstracting from 'temporary effects' such as the current recession and the war in the Persian Gulf, there is reason to believe that the longer term trend of U.S. fertility is upward, as a result of rising male relative incomes and female wages. Predicting the extent of such an upward trend, however, will require more effort in properly quantifying relative income and female wage effects.
Bibliography


unpubl. paper, Williams College Department of Economics, Williamstown, MA (December 9).


Willis, Robert J. (1987). "What Have We Learned From the Economics of the Family?" *AEA Papers and Proceedings* 7(2), 68-81.

A series of papers written by members of the Department of Economics on topics that do not pertain directly to less developed countries. From time to time documentation for computer programs and other research tools developed by members of the department will be distributed in this series as Special Research Papers and denoted by the letters RPS. A copy of any paper (and/or a reprint of the published version) will be mailed on request. See the order blank for further details.

1984-1993 Research Papers


RP-140 Roger Bolton, Randall Jackson, and Guy West, "Editors' Introduction" (to special issue of a journal, Socio-Economic Planning Sciences, vol. 23, No. 5,


** Reprint only available**
ORDER BLANK
FOR RESEARCH MEMORANDA AND RESEARCH PAPERS

Please send the indicated number of copies of each paper:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RM-69</td>
<td>RM-95</td>
<td>RM-121</td>
<td>RP-09</td>
<td>RP-35</td>
<td>RP-61</td>
<td>RP-87</td>
<td>RP-113</td>
<td>RP-139</td>
<td>RP-165</td>
</tr>
<tr>
<td>RM-77</td>
<td>RM-103</td>
<td>RM-129</td>
<td>RP-17</td>
<td>RP-43</td>
<td>RP-69</td>
<td>RP-95</td>
<td>RP-121</td>
<td>RP-147</td>
<td></td>
</tr>
<tr>
<td>RM-86</td>
<td>RM-112</td>
<td></td>
<td>RP-26</td>
<td>RP-52</td>
<td>RP-78</td>
<td>RP-104</td>
<td>RP-130</td>
<td>RP-156</td>
<td></td>
</tr>
</tbody>
</table>

My correct mailing address is:

MAIL THIS COMPLETED ORDER BLANK TO THE ADDRESS IN THE TOP LEFT CORNER.

* Please Note: There is a 10 paper limit with only one copy of each paper.