J Popul Econ (1998) 11:53–111

Journal of
Population
Economics

Fertility and the Easterlin hypothesis: An assessment of the literature

Diane J. Macunovich

Maxwell Center for Policy Research, 426 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020, USA
(e-mail: dmacunov@maxwell.syr.edu; Fax (315) 443-1081)

Received: 16 July 1996 / Accepted: 26 September 1997

Abstract. Focusing just on the fertility aspects of the Easterlin hypothesis, this paper offers a critical assessment – rather than just a selective citation – of the extensive fertility literature generated by Easterlin, and a complete inventory of data and methodologies in seventy-six published analyses. With an equal number of micro- and macro-level analyses using North American data (twenty-two), the "track record" of the hypothesis is the same in both venues, with fifteen providing significant support in each case. The literature suggests unequivocal support for the relativity of the income concept in fertility, but is less clear regarding the source(s) of differences in material aspirations, and suggests that the observed relationship between fertility and cohort size has varied across countries and time periods due to the effects of additional factors not included in most models.

JEL classification: J11, J13, N3

Key words: Easterlin, relative income, fertility

1. Introduction

Richard Easterlin generated considerable interest among many social scientists in 1978, when in presenting his inaugural presidential address to the Population Association of America he provided his answer to the question "What Will 1984 Be Like?" He based his predictions on changing birth co-

I would like to thank three anonymous referees and the editor for their valuable comments on previous drafts. *Responsible editor:* Klaus F. Zimmermann

hort sizes and their effects on what he termed 'male relative income', the income of young men relative to their material aspirations. Some of the ideas in that paper had been developed – and presented – earlier, even as early as 1961, but in bringing all of them together in this 1978 address he presented a highly persuasive explanation for the U.S. post-World War II baby boom. In addition, he used the baby boom to explain many of the puzzling phenomena which society had experienced since the 1960s – especially the baby bust – and offered some very attractive predictions for the 1980s.

Thus Easterlin's address was, to many, not only intellectually promising, it was also appealing at an intuitive level because of its optimism. But his paper attracted attention for another important reason: it challenged the orthodox neoclassical economic model of fertility originally suggested by Becker (1960), and then elaborated by Becker and Lewis (1973). Becker had developed a framework in which the demand for children could be treated as analogous to the demand for producer or consumer 'durable goods', depending on whether parents expected net pecuniary returns, or direct utility, from children, and fluctuations in fertility could be explained using only prices and incomes. Easterlin's model on the other hand, while still taking an economic approach, was based on the somewhat heretical notion of 'shifting preferences': preferences (i.e., material aspirations) which changed systematically as a function of the same variables used to predict fertility – income and prices. Becker's 'Chicago School' felt that there was no justification for assuming changes in preferences, and indeed felt that such an assumption was antithetical to the economic approach to fertility. But Easterlin and most sociologists saw the Easterlin model as an attempt to bring sociology into economics, and believed that an economics which couldn't accommodate changing preferences was not a feasible science.

1984 came and went, however, with almost no change in fertility rates. In addition, past trends appeared to continue unabated in the other indicators used by Easterlin: marriage and divorce rates, crime rates, female labor force participation rates, and male relative income. None of this helped in terms of economists' acceptance of the Easterlin model. Then a brief flurry of rising fertility rates in many developed countries in the late 1980s – accompanied by an apparent 'topping out' of younger women's labor force participation and divorce rates – tantalized supporters of the Easterlin hypothesis. Was it related to declining cohort size? What does a thorough review of the literature tell us: is further research on this topic justified?

Olsen's (1994) critical assessment of the leading contenders among fertility and female labor force participation models, including Easterlin's, concluded that "the recent stability of the fertility rate suggests these theories do not have good short term predictive power." While Olsen's critique was wide-ranging in terms of the number of models considered, a review by Pampel and Peters (1995) was wide-ranging in its attempt to cover the literature on all aspects of the Easterlin hypothesis. Their review is highly recommended, particularly because of its thoughtful discussions of potential mitigating effects in various contexts – for example, changing gender roles, the effects of immigration on labor supply, and the existence of societal institutions which cushion the harmful impact of large cohort size on economic well-being. It concluded that "the evidence for the Easter-

lin effect proves mixed at best and plain wrong at worst." Macunovich (1996b), however, draws on the fertility and female labor force participation literature to demonstrate that recent fertility experience is highly consistent with a synthesis of the Easterlin and New Home Economics models – as suggested by Sanderson (1976) in his thoughtful evaluation of these two models.

In addition, there have been selective reviews presented in an attempt to explain post World War II fertility patterns in developed countries, such as Sweezy (1971), Schultz (1981), Bean (1983), Sweet and Rindfuss (1983), and Ermisch (1982, 1983, 1990). These and others have tended to solidify the general impression that the Easterlin hypothesis 'works' at the macrolevel but receives little or no support in microlevel analyses.

But there has not, to date, been a thorough review which focuses just on the fertility aspects of Easterlin's hypothesis, and offers a critique – rather than just a selective citation – of the many studies contained in the literature. This paper attempts to offer a more critical assessment of the extensive fertility literature generated by Easterlin, with a complete inventory of data and methodologies employed. Many will be surprised – as I was – to find that a simple tally, when all studies are included, shows that there have been an equal number of micro- and macro-level analyses using North American data (twenty-two), and that the 'track record' of the hypothesis is the same in both venues, with fifteen of the twenty-two finding significant support in each case. In addition, an examination of the studies which provide least support might lead some to question whether they actually address the Easterlin hypothesis as he formulated it – but this could be said for several of the supportive studies as well! Sometimes because of data limitations, sometimes because of widely varying interpretations of the hypothesis, researchers have conducted studies which seem to bear little resemblance to the hypothesis. But this raises an important issue in reviewing and assessing the literature: what is the essence of the Easterlin hypothesis?

Is the primary significance of the Easterlin hypothesis its emphasis on *relative economic status*, and its identification of that concept with *the level of parental affluence* during an individual's formative teen years? Or is it in his application of that concept to the *post-World War II baby boom and bust*? Or in his – and others' – suggestion that such an effect could, in a fairly closed economic-demographic system, cause 'self-generating population cycles'? And to achieve significance must relative economic status be identified as *the sole cause* of the baby boom and bust?

The literature suggests unequivocal support for the relativity of the income concept in the fertility decision: it is essential to control for systematic differences in material aspirations across the population and over time. But the literature is less clear regarding the source(s) of differences in material aspirations – largely because studies have addressed such a wide range of potential sources.

And there is strong support for the role of systematically changing material aspirations in the post -World War II baby boom and bust – but with definite indications that this has been what Pampel (1993) calls a 'contingent' effect. Other factors have played a role, and the role of those other factors appears to have been significant in explaining differences across countries and time periods. But the jury is still out regarding 'self-generat-

ing cycles'. A great deal of work has been done to identify the necessary characteristics of models consistent with such cycles, and this work indicates an extremely small range of potential candidates (see the discussion of Wachter 1991 in Sect. 7).

This paper begins with a description of Easterlin's own tests of his hypothesis (Sect. 2), and then moves on to analyses conducted outside of North America (Sects. 3 and 4). United States and Canadian analyses are discussed in Sects. 5 and 6, followed by sections on population cycles (Sect. 7), preference formation (Sect. 8) and age-period-cohort analyses (Sect. 9). The paper concludes with a discussion of differences among studies, which appear to be related to differences in their findings on aspects of the Easterlin hypothesis.

2. Richard Easterlin on the Easterlin hypothesis

Before launching into a review of what *others* have said about the Easterlin hypothesis, we should trace the development of the idea through Easterlin's own work in order to understand the hypothesis as he saw it. We will find that to be a necessary backdrop as we begin to cover the wide range of interpretations that others have placed on the hypothesis. (Note that technical details of all the studies described here are summarized in Table 1.)

The earliest (published) forerunner of the hypothesis appeared in the *AER* in 1961, but at that time it was simply an attempt to explain post-World War II fertility in terms of relationships between *absolute* income and fertility behavior, with a strong emphasis on the fortunes and behavior of the young, primarily those in the age group 20–29 "where so many decisions regarding marriage and childbearing are concentrated" (p. 14). He emphasized population growth cycles, and traced the inverse relationship between the 1945–1960 pattern of fertility and the *rate of labor market entry* in any given year. Interestingly, the only reference to a 'relative cohort size' concept in this paper was his demonstration of a 'marriage squeeze' variable (an age-specific male:female ratio which falls when cohort size is increasing) operating on the marriage and fertility rates of foreign-born white women aged 20–29. In later papers, though, he found that the same data were fit much better using a relative income variable.

In this early paper he predicted an imminent precipitous decline in fertility rates. He suggested that "one might imagine" a "more or less self-generating mechanism" (p. 32), but said that this was just one hypothetical possibility. He stressed that the fundamental point was that economists and demographers should avoid thinking in terms of simple secular decline in the future, because "substantial fertility variation, up or down, may occur over the longer run" (p. 32).

By the time of his 1966 *Demography* article, Easterlin was writing from strength, since it was apparent that his prediction regarding fertility rates in the 1960s had come about, with a vengeance, and he predicted a continued decline through about 1970. He had arrived at and accepted the idea of shifting preferences, despite his biases as an economist against such a phenomenon, and began speaking of "the desired consumption level." Based

Table 1. Relative income and cohort size variables used by Richard Easterlin (all studies for the United States unless otherwise specified)

Study	Years covered	Relative income measure	Relative cohort size measure	Fertility measure	Other variables	Method of analysis
1961	1875–1929		'marriage squeeze'[a]	'fertility ratio'[b] (white foreign-born)		visual inspection
	1885–1958	real farm income		'fertility ratio'[b] (white rural)		visual inspection
	1885–1958	civilian unemployment rate	rate of change, # of white males 20–29	'fertility ratio'[b] (white urban and total)		visual inspection
1966 a	1870–1964		males 30–64/15–29	crude birth rate		visual inspection
1966 b	1953–1962	family income 14–24$_{t,\ 3\ year\ MA}$ / family income 45–54 $_{t\text{-}5,\ 3\ year\ MA}$		TFR and age specific rates 15–29		visual inspection
1968	as in 1966 a, b	as in 1966 a, b	as in 1966 a, b	as in 1966 a, b		visual inspection
1973 a	1947–1970	family income 14–24 $_{(t\text{-}1)\text{-}(t\text{-}5)}$ / family income 45–54$_{(t\text{-}2)\text{-}(t\text{-}8)}$ and $U_{(t\text{-}1)\text{-}(t\text{-}8)}/U_{(t\text{-}3)\text{-}(t\text{-}22)}$		married women 20–24		visual inspection
1976	1940–1975	as in 1973 a	males 35–64/15–34	TFR		visual inspection
1976 [e]	1940–1969/1971		males 35–64/15–34	TFR (four countries) [e]		visual inspection
1978 a, b [f]	1920–1975		males 30–64/15–29	crude birth rate		visual inspection
1978	1940–1977	as in 1973a	males 30–64/15–29	TFR		visual inspection
1980	1940–1977	as in 1973a	males 30–64/15–29	CBR w/RCS, TFR w/RY [g]		visual inspection
1982 [h]	1951–1976/1981		males 35–64/20–34	TFR (10 countries) [h]	none	OLS regression

[a] This is calculated as males aged 25–34 relative to females aged 20–29, in the foreign-born white population.
[b] This is the number of children under five, relative to the number of women 20–44, in each population.
[c] This is a three year moving average of the median total money income of all families with head aged 14–24 as a percentage of the (three year moving average) of the median income 5 years earlier, of families with head aged 35–44.
[d] This is the annual average total money income of families with head aged 14–24, e.g., averaged over 1953-1957, where the denominator would be the annual average total money income of families with head aged 45–54 averaged over 1950–1956. This estimated relative income would be applied to fertility in 1960.
[e] With Gretchen Condran. The countries included are the United States, England and Wales, Canada and Australia.
[f] With Michael and Susan Wachter.
[g] CBR = Crude Birth Rate, RCS = relative cohort size, RY = relative income.
[h] With Marc Artzrouni. The countries included are the United States, England and Wales, France, the Netherlands, Sweden, Denmark, Finland, Spain, Italy, Switzerland.

on the work of Duesenberry (1949) and Modigliani (1949) he expected that fertility would be inversely related to desired consumption, and credited Dorothy Brady and Rose Friedman (1947) for their "pioneering contribution on the 'relative income' hypothesis."

This article exhibited most of the elements of his fully developed hypothesis, especially his use of lagged parental income as a proxy for young peoples' consumption aspirations. He began here amassing the various time series he used over the next fifteen years to illustrate the workings of his hypothesis, including an estimate which showed the relative income of young males declining sharply from 1955–1957 onward, after having reached a ratio of about 4/5 in 1953. He developed his relative income figure here using a three year moving average of the total money income of families with head aged 14–24 as a percentage of the (three year moving average of) income five years earlier, of families with head aged 35–44. Although lacking the detailed data needed for this calculation prior to World War II, he developed a "rough impression" of the "differences among successive younger cohorts in inherited consumption desires" using net tangible assets of households from 1929–1958: the result was that "the cohorts reaching childbearing age, say 15–19, when asset levels were lowest, roughly in the decade 1940–1950, include those most instrumental in the baby boom" (p. 143). He emphasized that:

– income patterns of young people differed from those at the aggregate level: these discrepancies caused the observed changes in relative income, which in turn caused the observed changes in fertility rates;
– the fertility decline started first with 15–19-year olds in 1956–1958, with the other young age groups (20–24 and 25–29) not showing any tendency to decline until the period 1959–1961: a pattern consistent with the idea of relative cohort size effects;
– surveys of expected family size throughout the baby boom and bust periods showed no sign of change – but economic conditions when young caused cohorts to over-and under-shoot these expectations;
– the increase in female labor force participation rates in the 1960s differed greatly from that in the 1950s, when the increase was primarily among women in their 40s: in the 1960s it occurred mostly among mothers with young children, which he attributed to declining male relative income.

1969 saw the initiation of his emphasis on the need to use measures of lifetime 'potential' rather than current earnings in calculating male relative income – perhaps in response to the fact that his original measure of relative income appeared to diverge markedly from the pattern of relative cohort size in the mid 1960s. He began emphasizing changes in relative cohort quality as well as size, and in this article he presented his ideas regarding the relativity of childrearing prices as well as time costs, and the concept of regulation costs in fertility behavior.

By 1973 a divergence between his measure of relative income, and relative cohort size, became apparent: his measure appeared to be significantly off-track. Although in this paper (1973a) the numerator continued to be based on the income of families with head aged 14–24, the denominator in his relative income figure had become the income of families with head aged 45–54 (rather than the earlier 35–44), based on data showing the ac-

tual ages of parents of young people in the 14–24 range. He began here to calculate relative income by, e.g., dividing the 1953–1957 average income of the younger group, by the 1950–1956 average for the older group, and applying this figure to fertility in 1960. He pointed out that the relative income figure began to rise in the latter half of the 1960s while fertility leveled off, and stated that "there is some question as to the channels through which the influence of relative numbers is exerted" (p. 187). In this article he began to emphasize more the role of aggregate demand, over that of relative cohort size, and predicted that any future fluctuations in fertility will be "dampened in magnitude . . . because the biggest changes in relative economic status of adults . . . have been bound up with the Great Depression" (p. 213).

In the same paper, in order to look back to the prewar period, he developed an alternative measure of relative economic status using an eight year average of the general unemployment rate (assumed indicative of a young man's current and recent labor market experience) relative to a three year lag of a twenty year average of the same rate (representing the labor market experience of his father). The general unemployment rate was used because age-specific rates were not available back to 1930, *although this was seen as "a shortcoming of the approach"* (p. 194) since it would not capture effects related specifically to young adults. This relative status index proved to

> "accord reasonably well with the ups and downs shown by the fertility rate. There is some suggestion that the economic series leads the fertility series in timing in the post-World War II period, but in view of the much greater crudity of this relative status index than that used [previously], it is probably best not to make much of this timing difference" (p. 196).

It was in this 1973a article that he presented his 'cognitive dissonance' argument for preferring period to cohort measures of fertility: stated family size desires of a cohort will be influenced by current and past experience, which is reflected in period rates. He pointed out Blake's (1967) demonstration that the proportion of young couples desiring two to four children had been remarkably constant since 1936, between 85–95% – but the mean ideal family size had fluctuated widely within this range, as couples adjusted their ideal to their experience, given economic conditions. This percentage is echoed by the 88% found in a series of surveys of college students in the 1990s administered by the author (see for example, Goodwin 1990).

A number of critics of the hypothesis have focused on these seeming anomalies in income and wages in the late 1960s and early 1970s (see, for example, Smith 1981, 1986; Rutten and Higgs 1984; and related arguments for Great Britain in Ermisch, 1979, 1982 and 1983), but with the benefit of hindsight and considerably more data we now know that the relative income of young males in the United States rose only temporarily in the late 1960s as a result of the Vietnam War, and then resumed its precipitous decline as indicated in Fig. 1 (taken from Macunovich 1998). The pattern shown there for African-Americans and whites was mirrored in every education group, and at all points in the income distribution, and Macunovich

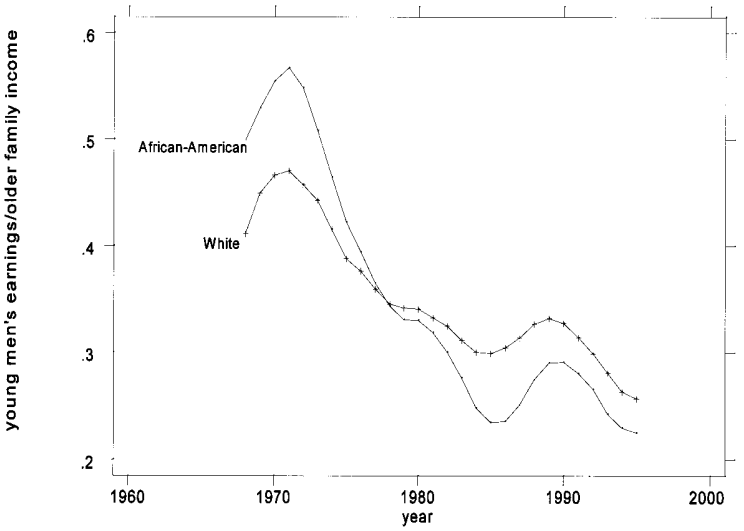

Fig. 1. Measures of African-American and white male relative income: average annual earnings of all men in the first five years of work experience as a proportion of the average annual income of families with head aged 45–54 five years earlier. (*Source:* Macunovich 1998)

(1997a) demonstrated that these patterns are a function of relative cohort size, augmented by the military draft and international trade effects. Similarly, although many arguments against the Easterlin hypothesis and in favor of the New Home Economics model of fertility allude to strong increases in young women's wages throughout this period, the only measure of female wages which rose throughout the 1970s was an endogenous one which incorporated the effects of increasing education and experience. When education and experience are held constant for young American women, as in Fig. 2 (taken from Macunovich 1998), it is clear that the only period of increase in their exogenous wage occurred in the late 1960s and early 1970s.

Consistent with these data, Easterlin's 1976 article demonstrated that the seeming anomaly in the figures in the 1960s appeared to have sorted itself out, indicating that the original relative cohort size argument still held. He continued using the same income measures as in his 1973a article: e.g., the 1953–1957 average income of families with head aged 14–24, relative to the 1950–1956 average for families with head aged 45–54, applied to fertility in 1960. But he added to that an age ratio – males aged 35– 64 relative to males aged 15–34 – developed in his 1976 study with Gretchen Condran (an article included in the discussion in Sect. 3).

In 1980 he presented a formal theoretical model of fertility choice with endogenous tastes (with Pollak and Wachter), and then in 1982 he presented another international comparison (with Artzrouni), this time using ten countries in the period 1951–1976 (included in the discussion in Sect. 3).

And finally, one last analysis by Easterlin which is directly relevant to the hypothesis is presented as an "Epilogue" in the second edition of *Birth*

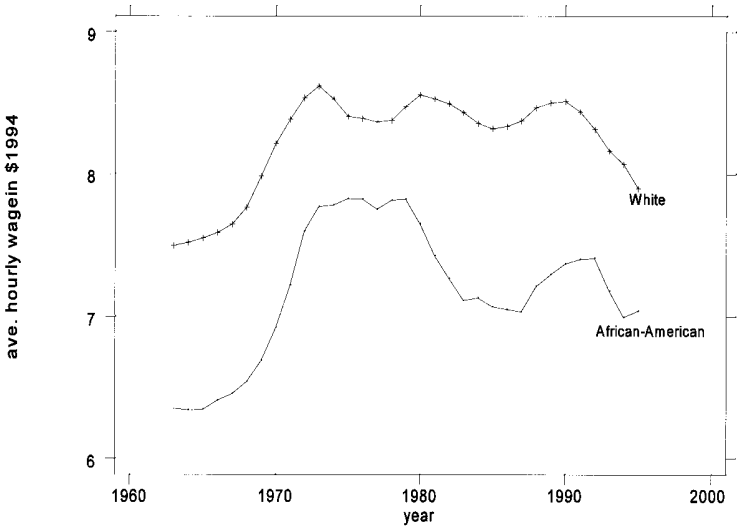

Fig. 2. Average hourly wage, in 1994 dollars, for African-American and white women in their first 5 years of work experience, holding education constant at the 1967 level. (*Source:* Macunovich 1998)

and Fortune (1987a), where he hypothesized that many of the "striking behavioral changes of the baby boom generation [primarily delayed marriage and family formation and increased female labor force participation] reflect chiefly the struggle of the baby boomers to maintain their economic status relative to their parents" (p. 165). Here, he looked at what he termed 'ex ante' and 'ex post' income for the baby boomers, where "the ex ante situation is the relative income status of young adults in the absence of any adjustments to changing market forces and the ex post situation, the relative income after such adjustments" (p. 166). His empirical analysis demonstrated an impressive performance of the baby boomers in managing to maintain their status relative to their parents. Despite a drop in male relative wages from 97.3% to only 79.7% between 1968 and 1982 (using males aged 25–34 relative to males aged 45–64), ex post relative income of young people (using family income on an 'income per adult equivalent' basis) improved from 95.7% to 99.0% during the same period.

Birth and Fortune and Easterlin's entries in the *New Palgrave* (1987b, 1987c) on "The Easterlin Hypothesis" and "Fertility" are essential reading for those searching for an explanation of Easterlin's ideas in his own words. Some related analyses include Crimmins et al. (1991); Easterlin (1995a); Easterlin et al. (1990a,b, 1993); Macunovich and Easterlin (1988, 1990).

3. Evidence from cross-country analyses

There have been seven multi-country studies using aggregate data to test Easterlin's hypothesis (in its relative cohort size formulation), which might be thought of as setting the stage for individual country analyses: Artzrouni

and Easterlin (1982); Baird (1987); Chesnais (1983); Easterlin and Condran (1976); O'Connell (1978); Pampel (1993); Wright (1989).

3.1 Summary of results

These seven are summarized in Table 2, with an attempt to compare their results, country by country. Such a comparison is fraught with hazards: it is based in some cases on empirical results presented by the authors, and – when these were not provided – on the authors' verbal assessments of their results. The seven studies were remarkably similar in that they all attempted to determine the relationship between relative cohort size and the TFR in each country, but their methodologies ranged from simple visual inspection to Granger tests of causality. Given the different methodologies employed, it is perhaps encouraging that there appear to be so many similarities in their results. For example:

– Australia, Belgium, Canada, England and Wales, New Zealand and the United States – in terms of strong supporting results;
– Germany, for evidence of an 'inverse' relationship (that is, higher fertility associated with larger cohorts); and
– Portugal and Spain, for the absence of any relationship.

Noting that a '1' in the table signifies a 'maybe' – the possibility of a relationship – we might also add Finland, France, and Norway into the first of these categories.

In the Netherlands, Austria and Japan there are not conflicting results (in the sense of researchers finding both positive and negative relationships) – just a few researchers who failed to find significance in the relationships. So perhaps we could add the Netherlands into the first category above, and Austria and Japan into the second (although Ohbuchi, 1982, suggests that relative income measures, rather than relative cohort size measures, are closely related to fertility in Japan).

Ireland is a puzzle, since only two researchers have studied it, with one producing evidence of a strong relationship and the other producing no statistical significance – both using the same time periods, variables and methodologies (Granger causality tests). For Italy and Sweden the majority of researchers found significant evidence of an Easterlinian relationship, but in each case there was one researcher who found evidence of an *inverse* relationship. And Denmark demonstrates no consistency whatsoever. The overall 'score' then is ten countries – possibly twelve, if we include Italy and Sweden – with supportive results, three with 'inverse' relationships, and four (Portugal, Spain, Denmark and Ireland) with no measured relationship.

3.2 Notable differences among these studies

Although there were strong similarities among the seven studies summarized in Table 2, there were a few significant differences which should be noted. Taking these chronologically as well as substantively, O'Connell's (1978) analysis differs most significantly from the others in his use of a unique cohort size indicator, which he intended to measure "cohort stresses": the ratio of the average number of males 20–34 in years *t-1* to *t-3*,

Table 2. A summary of results from seven cross-country relative cohort size analyses [a]

Authors	Artzrouni Easterlin	Baird [b]	Chesnais	Easterlin and Condran	O'Connell [c]	Pampel [d]	Wright
Date of study	1982	1987	1983	1976	1978	1993	1989
Approx. period covered	1951–1981	1950–1981	1930–1980	1940–1971	mixed [c]	1951–1986	1950–1985
Analytical technique	OLS/C–O	OLS/C–O	simple correlation	visual inspection	OLS	Granger causality	Granger causality
Cohort size ratio used	35–64/ 20–34	35–64/ 15–34 [b]	35–64/ 15–34	35–64/ 15–34	cohort position [c]	30–64/ 15–29	30–64/ 15–29
Australia			4	4	3	3	
Austria			–1		–4 [c]	0	0
Belgium			3		3	2	4
Canada			4	3	4	4	
Denmark	3		0		4	0	–1
England and Wales	4	3	3	3	4	3	4
Finland	4				3	1	2
France	3	4	3		0	1	3
Germany			–1			–1	–1
Greece							0
Hungary					0		
Ireland						4	0
Italy	–1		3			3	3
Japan			–1		–4 [c]	0	
Netherlands	4		0		3	4	0
New Zealand					3	2	
Norway			2		4	2	1
Portugal			1		0		0
Spain	2		1				1
Sweden	3	3	2		4	1	–1
Switzerland	4		–1		–4 [c]	–1	0
United States	4	4	4	4	4	4	

[a] All studies used TFR as the dependent variable and – except for Baird and O'Connell – conducted simple bivariate analyses of relative cohort size (RCS) and TFR. –1 indicates an inverted (i.e., 'wrongly signed' relationship, 0 indicates no measured relationship, 1 indicates 'maybe', and 2–4 indicate significant relationships with the 'correct' sign, with 4 assigned to the strongest. See footnote *c* for an explanation of –4.

[b] Baird used two cohort measures simultaneously – an age ratio and an unemployment ratio (the average of the last 5 years over the average of the previous 15 years) – together with the current unemployment rate.

[c] O'Connell used (males $20–34_{(t-1)–(t-3)}$)/(males $20–34_{(t-1)–(t-6)}$) as his cohort measure – a variable which would be larger for cohorts on the leading edge of the baby boom, and smaller for those on the trailing edge. Because of this difference, both positive and negative effects of his cohort variable are consistent with the Easterlin hypothesis. Japan, Austria and Switzerland all exhibited highly significant effects of cohort position, but opposed in sign to the effects in other countries – thus I have assigned them scores of –4. The years covered in his analysis vary over countries, from 1930–1975 (United States), to only 1956–1971 (Austria). He also included an unemployment distributed lag as a period control.

[d] Pampel included measures of 'collectivism' in his pooled analyses, but only the age ratio in his Granger causality tests.

over the average in the same age group in years *t-1* to *t-6*. Because of its construction, it would pick up differences between cohorts on the leading vs. lagging edge of a boom (since it would tend to be larger for the former, than for the latter), rather than differences between relatively larger and smaller cohorts. Thus, in his results a negative coefficient would indicate that cohorts born in a period of rising fertility suffer more than those born when birth rates are on the decline. And, because of the unique construction of his cohort effect variable, significant positive as well as negative coefficients would be consistent with the Easterlin hypothesis (although in need of additional theoretical interpretation).

Baird's paper differs from the others in his simultaneous use of two relative measures: an age ratio and a 'relative income' measure. The latter was not age specific, since it was constructed using two different lags of the general unemployment rate: the reader is referred to the discussion in Sect. 10.1 regarding shortcomings of this type of variable. Baird concluded, however, that his results supported the Easterlin hypothesis.

And finally, in addition to the results presented in Table 2, Pampel (1993) presented results incorporating a series of variables intended to control for institutional differences among countries, together with the female labor force participation rate (FLFP) in an attempt to incorporate Oppenheimer's (1976) ideas about the mediating effects of FLFP on relative cohort size effects (see the discussion in Sect. 10.4). He found that relative cohort size effects were strongest in those countries with the lowest 'collectivism' ratings, (especially Canada and the Unites States), and particularly in the 1950s and 1960s – but still measurable in a number of other countries, as well.

While his analysis is useful and informative, it raises many questions – especially with regard to his inclusion of the potentially endogenous female labor force participation rate. He acknowledged in a footnote the potential endogeneity of this variable in a fertility equation, and emphasized that the coefficients should be "treated with caution" , but he did not provide any regression results which excluded this variable.

4. Evidence from single-country studies outside of North America

Table 3 presents a summary of the countries and time periods covered, and variables used, in analyses focused on countries outside of North America. The dominant figure in these analyses has been John Ermisch, (and he has also been the most critical of the hypothesis), so this section of the literature review begins with his findings. His published work includes five empirical analyses – one co-authored (1988, with deCooman and Joshi), three alone using British data (1979, 1983, 1988), and one based on the German postwar experience (1980) – and two review articles which deal with the Easterlin hypothesis (1982,1990).

4.1 Analyses by John Ermisch

His two earliest studies (1979, 1980) attempted to reproduce relative income and cohort size measures used by Easterlin (1969) and Wachter

Table 3. Relative income and cohort size variables used in single-country studies outside of North America [a]

Study	years covered	Countries studied	Relative cohort size / income measure	Fertility measure	Other variables	Method of analysis
Bernhardt 1972	1954–1967 (micro)	Sweden	individual family income / average income in social stratum	fertility from 1954/58 marriage through 1967		bivariate analysis and graphical presentation
Carlson 1992	1947–1989	6 East European [b]	labor force (entrants-exits) [b] / total workforce	TFR		visual inspection
Congdon 1980	1946–1978	London, England	absolute size of all 5-year cohorts aged15–64	six 5-year age-specific fertility rates	variants included Fourier series terms	OLS regression
Danziger and Neuman 1989	1974 (micro)	713 Israeli Jewish families	own predicted wage / parents' socio-economic status	# of children and log-odds # >=5	six [c]	logit regression
de Beer 1991	1950–1985	Netherlands	own birth cohort size / birth cohort size 20years before	completed cohort fertility	none	OLS regression
deCooman Ermisch and Joshi 1988	1971–1985	Great Britain	cohort size$_{\text{birth cohort at age 14}}$ / cohort size$_{\text{largest cohort, born in 1948}}$	age- and order-specific fertility and completed family size	six [e]	OLS regression
Ermisch 1979 **	1955–1975	England and Wales	males 35–64/15–34	TFR		visual inspection
			expected [d] wage$_{\text{all males in t-(t-6)}}$ / expected wage$_{\text{all males in (t-6)–(t-15)}}$	fertility of married women 20–24		visual inspection
			expected wage$_{\text{all males in year t}}$ / expected wage$_{\text{all males, 10 year MA}}$	TFR		visual inspection
			$U_{(t-1)-(t-8)}$ / $U_{(t-6)-(t-15)}$	TFR		visual inspection
			$U_{(t-1)-(t-8)}$ / $U_{(t-6)-(t-25)}$	TFR		visual inspection
			$U_{\text{3 year MA}}$ / $U_{(t-1)-(t-10)}$	TFR		visual inspection
Ermisch 1980 **	1935–1942 +1958–1977	Germany	expected [d] wage$_{\text{all males in year t}}$ / expected wage$_{\text{all males, 7 year MA}}$	TFR	women's real wage, % men in agriculture, + lags	OLS regression, Cochrane-Orcutt, first differences
			per capita GNP $_{t}$ / per capita GNP $_{\text{7 year MA}}$	TFR	women's real wage, % men in agriculture, + lags	OLS regression Cochrane-Orcutt, first differences

Table 3 (continued)

Study	years covered	Countries studied	Relative cohort size / income measure	Fertility measure	Other variables	Method of analysis
Ermisch 1983	1950–1977	Great Britain	men 30–64/15–29	TFR	none	visual inspection
			CBR in cohort year of birth	completed cohort fertility by 30 or 35	none	visual inspection
			men 30–64/15–29	age-specific fertility, 20– 24, 25–29, 30–34	male income, female earnings and LFPR	GLS regression
Ermisch 1988	1971–1985	Great Britain	cohort size$_{\text{birth cohort at age 14}}$/ cohort size$_{\text{largest cohort, born in 1948}}$	age- and order- specific fertility and completed family size	six[f]	tests for cointegration, OLS regression (dynamic specification)
Ohbuchi 1982	1947/ 1954–1980	Japan	men 35–64/15–34, average wage$_{\text{all males in (t-1)–(t-3)}}$/ average wage$_{\text{all males in (t-3)–(t-7)}}$, expected wage$_{\text{all males in t}}$/ expected wage$_{\text{all males in t-(t-7)}}$, average wage $_{\text{t-(t-2)}}$/ (family wage[g])$_{\text{(t-2)–(t-6)}}$	TFR		visual inspection
Serow 1980	1950–1976	Netherlands	average male wage$_{\text{(t-1)–(t-3)}}$/ average male wage$_{\text{(t-1)–(t-15)}}$ and size of marital cohort/ population 20–34	# children born, by marriage cohort	women's wages and LFPR	OLS regression, levels and differences

** Indicates a study which provided little or no support for the Easterlin hypothesis.

[a] Includes East- and West- Europe, Japan, Australia, and Israel. All analyses make use of macro-level data unless otherwise specified.

[b] Czech Republic, Slovak Republic, Bulgaria, Hungary, Poland, Romania. Labor force (entrants-exits) calculated as: population aged 15–24 minus (males 60–64 + females 55– 59).

[c] Mother's education, mother's education squared, father's education, mother's siblings, father's siblings, duration of marriage.

[d] The 'expected' wage is the average real wage of all males, multiplied by their employment rate. Subscript notation such as '(t-6)–(t-15)' is meant to describe the average of annual levels in each year of the designated period, e.g., from year *t-15* to year *t-6*.

[e] Men's hourly wage, male unemployment rate, rate of change in price level, time trend, lifetime employment rate of cohort, share of cohort at risk of j*th*-order birth.

[f] Female/male wage ratio, male net weekly earnings, male unemployment rate, rate of change in price level, real house prices, additional child allowance for j*th* order child, share of cohort at risk of j*th*-order birth

[g] 'family wage' is constructed as [male wage + (FLFP*female wage)].

(1975). The first consisted entirely of visual analyses of graphically presented data for England and Wales, while the second carried forward this type of analysis for Germany using OLS regressions.

His 1979 article began by examining the relationship between the TFR and a relative cohort size variable which he felt showed a fairly high positive correlation. He then went on to construct 'relative income' variables. However, as can be seen in Table 3, none of these 'income' measures were based on *age-specific rates* depicting the relationship between the incomes of older and younger workers. Instead, because of data constraints all approximated what he termed 'hypothetical' sons:fathers ratios using two different lags of the same aggregate time series – aggregate male wages, the aggregate unemployment rate, and per capita GNP – and he saw little correspondence between these variables and fertility.

In concluding this visual analysis Ermisch seemed to hold out some hope for the hypothesis: he stated that although relative economic status might not be the dominant influence upon fertility suggested by Easterlin, it might still be one of a group of influential factors. However, no attempt was made in that study to control for any other factors, and the author's general conclusion was to reject the Easterlin hypothesis *to the extent that his chosen non-age-specific relative economic status measures reflected the spirit of that hypothesis.* His 1980 analysis of German data used OLS to control for several other variables in addition to 'relative income' (calculated again as a ratio of two different lags of the average expected wage for all males), and still arrived at a negative conclusion – but here again based on an analysis using only non-age-specific measures of relative income.

Ermisch's 1982 and 1983 review articles concluded, on the basis of these two studies of his own, as well as MacDonald and Rindfuss (1978), Crimmins-Gardner and Ewer (1978), Olneck and Wolf (1978) and Thornton (1978) (all discussed in Sect. 6.1), that the relative income model was only weakly supported by the evidence.

In 1983 Ermisch returned to the use of relative cohort size measures, which he included in a model of age-specific fertility rates based on Butz and Ward (1979) for Britain in the period 1950-1977. Using these measures he concluded that relative generation size has a significant effect only on the fertility of women in their 20s. This is, of course, the age group on which Easterlin focused.

He continued in the 1980s with a relative cohort size measure, including it in dynamic models of sequential fertility decision-making, again for Great Britain, using data for 1971–1985. Here again, however, his chosen measure was not 'relative' in the Easterlinian sense. That is, rather than representing the size of a young cohort relative to that of its parental cohort, his measure was calculated as the size of an individual's cohort at age 14 relative to the size of the largest (1948) birth cohort – thus proxying only very indirectly the imperfect substitutability between older and younger workers. Despite this limitation, however, in a co-authored piece in 1988 (with de Cooman and Joshi) it was concluded that women from larger generations are less likely to start childbearing in their early 20s, and that they are also less likely to proceed from two to three offspring before they are 30.

In a sole-authored article also dated 1988, Ermisch estimated what he considered to be a more robust dynamic specification which included more

variables and dealt more rigorously with nonstationarity – but using the same relative cohort size measure as in his work with de Cooman and Joshi. He concluded in this article that

> "[a]ll else equal, women from larger generations appear to be less likely to begin childbearing in their twenties, ending up with moderately fewer children on average. This is consistent with Easterlin's hypothesis that large generations may have dimmer economic prospects, which causes them to postpone childbearing and have smaller families" (p. 571).

Ermisch presented the results of simulations (p. 574) in which he varied one factor at a time, producing the following effects on an average family size of 1.98:

+35% in women's hourly relative to men's weekly earnings	–0.30 child
+100% in the real value of the child allowance	+0.17 "
+100% in real house prices	–0.16 "
+ 20% in relative generation size	–0.13 "

It is not clear from the article why much larger changes were simulated in variables other than relative generation size, but it seems from the above figures that one of the strongest elasticities of completed family size was with respect to relative generation size. In addition, aspects of the Easterlin hypothesis suggest that variables such as women's relative wages and real house prices might well be endogenous, affected in turn by relative cohort size (Fair and Dominguez 1991; Fair and Macunovich 1996; Mankiw and Weil 1989). Despite these fairly positive results in his 1983 and 1988 studies, however, in his latest review article, based largely on this work and that of Robert Wright (1989), he came down fairly unambiguously against the Easterlin hypothesis.

4.2 Tests by researchers other than Ermisch

There are four additional tests of the Easterlin hypothesis based on European macro-level data, (Congdon 1980; de Beer 1991; Carlson 1992; Serow 1980) and a micro-level analysis (Bernhardt 1972), plus a micro-level analysis of Israeli Jewish families (Danziger and Neuman 1989) and a macro level analysis for Japan (Ohbuchi 1982). The authors of all of these studies felt that their results supported the relative income hypothesis, although deBeer was somewhat grudging in his acknowledgement. The following discussion highlights some of the most notable points regarding this set of studies.

Carlson (1992) presented a novel variation on the Easterlin model – "conditioned on Kornai (1982)" – in his analysis of six Eastern European countries, where he said that "we must think of Easterlin standing on his head." He explains that because of the soft budget constraint in socialist countries,

> "the larger supply of labor represented by the large cohort meets the insatiable appetite of monopolistic state firms for labor. Everyone is employed. Attention must then focus on the total volume of the wage bill

> rather than on the level of individual wages. This total volume expands, creating more demand for the products of the system. Additional workers reduce bottlenecks and other inefficiencies in the production stream. The result is a mild economic 'boom' during which it is easier to satisfy the queued-up demand for all sorts of system outcomes (housing, consumer durables, other consumer goods, education, and so on). The population responds in classic Easterlin fashion with higher fertility. . . On the other hand, when a large cohort is followed some years later by an unusually small cohort, all of these effects are reversed" (Carlson 1992:673).

Carlson found very little support in Poland and Slovakia, but for the other four countries he felt that empirical trends seemed to support the hypothesis derived from Kornai and Easterlin.

Ohbuchi (1982) examined the visual correlation between the TFR for Japan and a) a relative cohort size variable and b) three different variants of the relative income variable (based on Easterlin 1973; Michael Wachter 1975; Oppenheimer 1976). He found a very bad fit after 1960 for the relative cohort size variable, but a fairly good match using all of the three relative income variables – especially the 'Oppenheimer variant' (see Sect. 10.4 for a discussion of Oppenheimer).

Perhaps the weakest results in this set appear in Danziger and Neuman's (1989) analysis of Israeli microdata, which was hampered by fairly crude categorizations of parental status. Both the wife's and the husband's fathers were classified as 'low' or 'high' occupational status, and having 'attended' or 'not attended' school, with a third indicator for origin in Africa or Asia. No information was available for the couple's mothers. These categorical variables were then used as controls in a regression of number of children born on wife's and husband's (predicted) wage rate and education, number of siblings, and duration of marriage. Their results are somewhat difficult to interpret, because of the instability of coefficients over different specifications and what appears to be a multicollinearity problem between the couple's two fathers. However, there does seem to be a strong positive effect of having a father with low educational status, which the authors interpreted as support for the hypothesis.

5. Tests based on U.S. and Canadian aggregate time-series data

All of the studies discussed in this section are summarized in Tables 4 (Canadian) and 5 (U.S. aggregate analyses).

5.1 Unsupportive aggregate analyses using U.S. data

The bulk of these analyses are attempts to demonstrate that – within an age-period-cohort framework – age and period are sufficient to explain nearly all of the variance in the time series of age specific fertility rates in the United States (These age-period-cohort studies will be discussed together in Sect. 9.) The only studies which fall outside of this category are DeFronzo (1976) and Rutten and Higgs (1984), with the latter really a refu-

Table 4. Relative income and cohort size variables used in analyses of Canadian data

Study	Years covered	Relative income measure	Relative cohort size measure	Fertility measure	Other variables	Method of analysis
Abeysinghe 1991	1921–1988		population 30–64/15–29	TFR and age specific for women 15–29	none	cointegration tests and regression procedures[a]
Abeysinghe 1993	1951–1980/1986	income of male family heads <25 and <35/45–54		TFR, completed fertility and age specific 15– 24, 25–34	average weekly salary, female clerks	cointegration tests
Ahlburg 1986	1941–1979		f (Births $_{t-1}$, Births $_{t-2}$)[b]	total live births in year *t*	immigration correction	maximum likelihood correcting for serial correlation
Chaudhury 1977	1967–1968 (micro)	actual income/ predicted income[c], categorized as high, medium or low		length of first, second and third birth intervals, and fertility given marriage duration	ten[d]	multiple classification analysis and ANOVA
Wright and Maxim 1987[**]	1922–1984		population 30–64/15–29	TFR		visual inspection
	1921–1985		women 30–64/15–29	seven 5-year age-specific fertility rates	age and period measures	age-period-cohort analysis

[**] Indicates a study which provided little or no support for the Easterlin hypothesis.

[a] "The results . . . were derived using a procedure called fully modified (FM) estimation procedure for cointegrated regressions (p. 194)."

[b] He tested two sets of lags as *t-1* and *t-2*: 20 and 40 years, and 25 and 50 years.

[c] Chaudhury constructed a 'present' (husband's actual income in 1967 relative to that predicted for him on the basis of occupation, education, age and place of birth) and a 'past' (husband's income at marriage relative to income predicted on the basis of age at marriage, occupation at marriage, place of birth and education at time of survey) relative income measure.

[d] "husband's income . . . wife's education, wife's religion, husband's occupation, husband's education, duration of marriage, wife's age at marriage, current family planning status (p.182)." In some formulations wife's current and lifetime labor force status were also used, but Chaudhury excluded these variables in others, acknowledging possible endogeneity.

tation of Easterlin's positive findings, rather than an original empirical analysis. Some might want to include Butz and Ward (1977, 1979) here, as well – but my review of those studies leaves me undecided about the strength of their evidence against the Easterlin hypothesis.

Butz and Ward provided us with probably some of the most widely-known tests of the relative income hypothesis, by including three different formulations of the relative income variable in equations with the female hourly wage, with and without absolute male annual earnings. Only two of these formulations were carried over from their 1977 report to their published 1979 paper: one based on Easterlin (1973a) which uses a ratio of two different lags of the general unemployment rate, and another based on Michael Wachter (1975) which uses a ratio of two lags of the average wage – formulations Easterlin had felt were less than ideal because they do not specifically measure the experience of young adults. Neither of these two performed well when used with just the female wage, producing insignificant coefficients and R^2 of only 0.74 and 0.76. When male annual earnings were added, however, the Wachter-type variable was positive and significant with an elasticity of 1.306 as compared to an elasticity of only 0.691 for (positive and significant) male annual earnings.

The formulation which never made it to publication was a more sophisticated one (again based on Easterlin 1973a) which compared young family incomes with (lagged) older family incomes, but covered only the years 1957–1974 because of data limitations. This relative income variable was positive and significant when used with just the (negative and significant) female wage, with an *elasticity of 2.06 and an* R^2 *of 0.86*. Interestingly, absolute male annual earnings produced a *negative* coefficient, and the female wage a *positive and insignificant* coefficient, for all regressions in which those two were used together (without a relative income measure), in this 1957–1974 time series (1977:22, Table 4).

Rutten and Higgs (1984) conducted a visual analysis using Easterlin's own data, showing that with different scales, and different time periods for the variables, the apparent relationship between fertility and relative income looked considerably less convincing. However, there is a tendency in this article as well as in Smith (1981, discussed in more detail as an age-period-cohort analysis) to take points out of context and set up 'straw men'. An example of this type of argument is the statement in Rutten and Higgs that Ben-Porath (1975) is a microlevel analysis which does not support the Easterlin hypothesis. However, their reference (1984:208) is to only one column in one of seven tables of results in Ben-Porath's paper, which displays the only negative but insignificant coefficient of first-generation predicted earnings on second-generation fertility. What Rutten and Higgs failed to point out is that this insignificant result occurred when Ben-Porath introduced (into an equation with a significant negative coefficient for first generation earnings, which supported the hypothesis) another variable (farm background) which was highly correlated with the first.

An example in Smith is his use of an unpublished (and undated) paper by Lee, rather than the several articles published by that researcher in the 1970s, and his implication that Lee had associated deleterious cohort size effects with *any* increase in population size: "In modern economics it is not believed that the populations of developed countries are made poorer by in-

creases in their numbers, and few economists would agree with Lee that this is suggested by the patterns of U.S. fertility and economic growth between 1917 and 1972" (pp. 248–250). But in fact Lee had associated such effects only with increases in excess of an exponential time trend.

DeFronzo (1976) made use of an interesting methodology which, like Lindert's (1978), exploited potential earnings variation by state to conduct the usual macro-level analysis (using current income of young men relative to lagged income of older men in their states of origin) at a somewhat more disaggregated level, using 1950, 1960 and 1970 census data. DeFronzo found a positive and significant effect of absolute male income, but no significant effect of relative income on fertility. (Lindert, who used decennial census data from 1900 through 1970, found significant effects using pooled data for the entire period, for the post-World War II period, and for most single years on their own, but also found no significance in 1960 and 1970. He attributed this to a marked lack of variance in incomes by state in these later periods.)

It should be noted that DeFronzo's is one of several studies (Crimmins-Gardner and Ewer 1978, is another) which entered both *absolute* and *relative* income measures simultaneously. (Lindert used 'younger' and 'older' male incomes as two separate variables.) Easterlin's hypothesis is that relative *instead of* absolute income should be used in the analysis of fertility decisions. This is accomplished either by entering the husband's income and the selected 'denominator' as separate variables, or as a ratio (if their coefficients prove to be statistically indistinguishable). It is difficult to interpret an equation in which *both* are entered. Although DeFronzo reported regressions in which absolute male income was entered on its own, he did not report any in which relative income was entered on its own.

5.2 Supporting evidence from analyses of aggregate North American data

In addition to Easterlin's own analyses, there have been fifteen studies which fall into this category (sixteen, if we include the strong performance of Easterlin's variable in Butz and Ward 1977): Abeysinghe (1991,1993); Ahlburg (1982,1983,1986); Devaney (1983), Evans (1990); Lee (1976); Lindert (1978); Macunovich (1996a, 1998); Moffitt (1982); Schapiro (1988); Shields and Tracy (1986); Michael Wachter (1975). Unlike the remainder of these studies, Abeysinghe's analyses and Ahlburg (1986) looked at Canadian cohort size/relative income effects, so our discussion begins with them.

Abeysinghe (1991) used cointegration techniques to test for an equilibrium relationship between one of the standard relative cohort size (RCS) measures, (population 30–64/15–29), and both the TFR and a 15-year age-specific rate for women aged 15–29 – which was selected because it was more consistent with the definition of RCS. He found a close co-movement between 1940 and 1976 but a marked departure thereafter, and he attributed this inconclusive (or negative) result to the lack of fit between RCS and male relative income in Canada. He emphasizes that use of only RCS does not provide an adequate test of the Easterlin hypothesis, and that there appears to be a significant relationship between fertility and relative income

in the period since 1976, even though there is no apparent relationship between fertility and RCS.

He tested this relative income/fertility relationship in his 1993 article, where he proposed to combine the New Home Economics and relative income models by regressing fertility on the income of younger male family heads, older male family heads (45–54), and the female weekly wage (which he approximated using the weekly salary of female clerks). Adjusting for cointegration effects and multicollinearity, he found that the effect of parental income on fertility and female labor force participation was much stronger than that of own income. His five equations (for 15–24 and 25–34 age-specific rates, completed fertility, TFR and LFPR) for the period 1951–1988 all have R^2 over 0.90, and he found that for completed fertility the relative income elasticity exceeded the wage elasticity, whereas the opposite was true for TFR and age-specific rates. Contrary to the findings of Chaudhury (1977) – reviewed in Sect. 6.2 – he interpreted this as indicating that relative income has its strongest effect on completed fertility, and the female wage on timing. However, it should be noted that a number of adjustments were made to correct for multicollinearity, which may have affected the relative magnitudes of the coefficients on these two variables. A timing effect of relative income (RY) seemed to show up in a reversed sign on the coefficient of RY, for fertility among 25–34 year olds, suggesting that they had postponed births due to low RY at earlier ages.

Using Canadian data for the years 1941 through 1979, Ahlburg (1986) implemented Samuelson's (1976) formulation of Easterlin's model, in which births in the current year are a function only of births with two lags – either 20 and 40 years, or 25 and 50 years. He found that all coefficients were significant and correctly signed, and concluded that the relative cohort model was a good representation of the variation in post-war total live births.

In his 1982 analysis of U.S. data, Ahlburg tested not only the relationship between relative cohort size (RCS) and fertility, but also (in the 1953–1976 period) that between RCS and male relative income, and relative income and fertility. He found significant coefficients with expected signs in the 1921–1976 and 1941–1976 and 1953–1976 subperiods, but not in the entire 1904–1976 period, a difference which he attributed to the changing role of immigration in labor supply. In all cases he was able to explain well over 90% of the variance – in many cases well over 95%. He used his results to posit and test a "closed demographic model" which could produce "self-generating cycles of births". Then, in his 1983 paper he tested a formulation of this closed demographic model derived by Samuelson (1976), in which current live births are a function only of births twenty and forty years earlier. He felt that his simple version of the closed model derived by Samuelson not only tracked the United States experience well but also predicted a revival of fertility.

Four other early studies, along with Ahlburg's results, gave the Easterlin hypothesis its reputation for success with macrolevel data: Michael Wachter (1975); Lee (1976); Lindert (1978); Moffitt (1982). Wachter (1975) provided what many considered to be a highly convincing set of macro results for predicting GFR (total and by race), CBR and TFR (pointing out that his results were strongest for the most refined measure, the TFR). His rela-

tive income variable has been emulated by several others: a ratio of the expected wage in year *t* to a 10-year moving average, ending in *t*, of the same expected wage – where the expected wage is the total male wage multiplied by the employment rate. Along with the relative and absolute income measures, he included a dummy variable for World War II, since his data series began in 1925, the first lag of the dependent variable and a measure of urbanization – as well as an interaction term between relative income and urbanization. His estimated coefficients were all significant with the expected signs, and the model produced R^2s of around 0.95. The visual fit of his predicted and actual results is very impressive, over the entire period. The strongest criticisms of his model center on the interpretation of the urbanization variable (which he saw as a proxy for the relative price of a child) and its interaction, and the dependence of his very strong results on the lagged dependent variable. It should be noted, too, that his 'relative income' variable is not age-specific, and therefore fails to capture much of the cohort size effect focused on by Easterlin's hypothesis.

Lee (1976) constructed two relative cohort size measures using proportional deviations from exponential time trends – one based on birth cohort size and the other on the size of the total labor force aged 15–64 – hypothesizing that completed cohort fertility should be a function of the first, and period total fertility rates a function of the second. The second measure differed from Easterlin in that it assumed that when the size of the total labor force is large relative to trend, the incomes of all workers would be depressed below their long-run trend, resulting in lower fertility. He found that period TFR elasticity with respect to this second measure was –7.5, while completed cohort fertility elasticity was only about –1.0 with respect to the cohort size measure. He suggested that these two models were too inflexible in their complete reliance either on period or cohort effects, and indicated that other results showed that in reality both had been operating simultaneously in the Unites States.

Lindert (1978) made ingenious use of census data which enabled him to conduct both cross-sectional and pooled time series analyses of U.S. data from 1900–1970. His measure of relative income was based on the idea that an individual's prior economic history (and hence material aspirations) could be proxied using his state of origin, capitalizing on the variation across states in per capita incomes. In a series of regressions which included a number of other controls (see Table 5), one of which was the relative cost of children, he found that the past income variable always affected fertility negatively and usually passed the standard tests of statistical significance. He concluded that income history was an important determinant of the material standards that make couples limit family size. In his regressions for the postwar period he was able to explain 85–90% of variance in the child-woman ratio, and in the pooled regression for the entire 1900–1970 period his R^2 was 0.765.

Lindert's current income variable performed much less consistently, but he pointed out that its use assumed that the current personal income per worker for all age groups was a fair reflection of the income prospects of young adults. He explained the three cases where his past income variable was only weakly significant (1900, 1960 and 1970) as resulting from an abnormally high correlation between past and current incomes in the 1900

Table 5. Relative income and cohort size variables in United States aggregate analyses

Study	Years covered	Relative income measure	Relative cohort size measure	Fertility measure	Other variables	Method of analysis
Ahlburg 1982	1904–1976	average annual earnings in mfg. $_{t}$ / [a] average annual earnings in mfg. $_{(t-1)-(t-10)}$	males 16–29/30–64, and CBR $_{t-20}$ / CBR $_{t-40}$	crude birth rate	none [a]	OLS regression
Ahlburg 1983	1949–1978		f(Births $_{t-20}$,Births $_{t-40}$) [b]	total live births	none	OLS regression w/ Cochrane-Orcutt
Butz and Ward 1977	1957–1974	family income 14–24 $_{(t-1)-(t-5)}$ / family income 45–54$_{(t-2)-(t-8)}$		fertility of women 20–24	female hourly wage, w/and w/out male ann.earnings	OLS regression
Butz and Ward 1979**	1948–1974	$U_{(t-1)-(t-8)}$ / $U_{(t-3)-(t-22)}$ and expected average wage $_{t}$ / expected average wage $_{(t-1)-(t-10)}$		fertility of women 20–24	female hourly wage, w/and w/out male annual earnings	OLS regression
DeFronzo 1976**	1950, 1960, 1970 (state-level)	median income males 20–24$_{t}$/ median income males 35–44$_{t-10}$		children ever born to women 20–24$_{t}$	median income, males 20–24$_{t}$ and six others [c]	OLS regression
Devaney 1983	1947–1977	average cohort deviation from trend rate of increase in age-specific male income		desired completed fertility	predicted female wage	simultaneous equations: fertility and FLFP
Evans 1990	cohorts born 1905–1954		% deviation from trend in cohort size at 1 year of age	median age at 1st birth, mean and SD of age at fertility, % childless, mean completed fertility	unemployment, FLFP and economic growth at age 20–24; % HS grads; % college grads; %foreign born; farm origin; contraceptive knowledge	OLS regression with Cochrane-Orcutt correction
Kelly and Cutright [d] 1984**	1917–1976		none: [d] age and period only	age and parity-specific fertility rates	marriage, unemployment, pill-IUD, sterlization, year	age-period-cohort analysis
Lee 1976	1917–1974		trend residuals of log(pop'n 15–64), and trend residuals of log(birth cohort size)	TFR cohort completed fertility		OLS regression on time trend, plots of residuals against fertility measure

Table 5 (continued)

Study	Years covered	Relative income measure	Relative cohort size measure	Fertility measure	Other variables	Method of analysis
Lindert 1978	1900–1970 (state level)	income per worker $_{t}$/ state per capita income $_{t-20}$		child-woman ratio, 0–4/15–49	eleven[e]	OLS regression
Macunovich 1996, 1998	1969–1994	all and by race: (expected earnings, males in first 5 years work experience)$_{MA5,t-1}$/ (income of families w/head 45– 54 of either sex)$_{MA5,\ t-6}$		fertility of women 20–24, all and by race	female wage and unemployment (all and by race)	logistic regression
Moffitt 1982	1947–1975	predicted lifetime wealth cohort age a_t/cohort age a+25$_t$		5 age-specific rates 20–44	age, time, own-wealth, parental-wealth, and age-wealth interactions	GLS and OLS regression
Rindfuss, Morgan and Swicegood 1988[**]	1913–1978	none: period variables and dummies used instead		first-birth probability at ages 25, 30, 35	nine[f]	OLS regression
Rutten and Higgs 1984 [**]	1940–1975	family income 14–24 $_{(t-1)–(t-5)}$/ family income 45–54$_{(t-2)–(t-8)}$ and $U_{(t-1)–(t-8)}$ / $U_{(t-3)–(t-22)}$		TFR		visual inspection
Schapiro 1988	1954–1985	expected[g] male fulltime earnings 20–24/25–64, expected[g] male fulltime earnings 25–34/35–64	males 20–24/25–64, males 25–34/35–64	fertility rate 20–24, fertility rate 25–34	age specific female income, participation rate, unemployment rate	nonlinear 3SLS: fertility, FLFP, marriage, divorce
Smith 1981 [**]	1947–1974	none: age-period-cohort analysis	none	five year age-specific rates	none	age-period cohort analysis
Shields and Tracy 1986	1920–1980		population 18–24/25–64	TFR	FLFP, personal income, infant mortality	OLS regression
	1946–1980		population 18–24/25–64	TFR	as above +housing costs	OLS regression

M. Wachter 1975	1925–1968/1972	expected[g] wage $_{\text{all males in year t}}$/ expected[g] wage $_{\text{all males, 10 year MA}}$	CBR, GFR, TFR, GFR_{white}, $GFR_{nonwhite}$	absolute expected wage, urbanization, World War II, RY*urbanization	OLS regression, lagged dependent variable

[**] Indicates a study which provided little or no support for the Easterlin hypothesis

[a] Ahlburg regressed both his relative income figure and his fertility measure separately on each of the two relative cohort size measures. Subscript notation such as '*(t-1)–(t-10)*' is meant to indicate an average of the annual values in the indicated time period.

[b] Here Ahlburg tests a formulation of the variable based on Samuelson (1976).

[c] DeFronzo included *absolute* younger male income along with *relative* (=absolute younger male income / older male income), as well as female income, female education, % Black, % Catholic, urbanization, and land area per person in each state.

[d] Kelly and Cutright (1984) contained no cohort analysis. It is included here because it was cited in Pampel and Peters (1995) as having "found no effect of relative cohort size on yearly changes in fertility in the United States". This study found strong effects of a 'marriage' variable throughout, which could have been proxying for cohort size effects.

[e] In his pooled cross-section time-series regressions of state-level data 1900–1970, Lindert included: share of women 15–29 ever-married; share of women 30–44 ever married; share of women 15–49 who are 20–34; share of women 15–49 who are non-white; percent foreign-born; percent of males employed in agriculture; three regional dummies, South, New England, and West North Central plus Mountain; and dummies for 1960 and 1970. In some formulations he also included the relative cost of children and each state's child-woman ratio twenty years earlier.

[f] Dummy variables for World War I, World War II, and the post-World War II period; military enlistment between 1942–1945; the unemployment rate between 1928–1945 (representing the Great Depression), and again in the period after 1945; the inflation rate in the period after 1945; and two 'CHANGE' variables – the first set to one between 1946–1950 for those aged 25, and between 1946–1953 for the other two age groups, and the second set to one between 1954–1955 for those aged 30 and between 1954–1960 for those aged 35.

[g] The 'expected' wage is the average real wage of all males, multiplied by their employment rate. Subscript notation such as '(t-6)–(t-15)' is meant to describe the average of annual levels in each year of the designated period, e.g., from year *t-15* to year *t-6*.

Table 6. Relative Income and cohort size variables in United States microlevel analyses

Study	Years covered	Relative income measure	Relative cohort size measure	Fertility measure	Other variables	Method of analysis
Bahr, Chadwick and Stauss 1975	1973 (micro)[a]	own occupation/ father's occupation, current occupation/ desired occupation, own lifestyle/ parents' lifestyle at same age, own income/ income in same age-occupation-education group; own lifestyle/lifestyle in same age-occupation-education group		number of children	age, marriage duration, number of siblings	stepwise OLS
Bean and Swicegood 1979	1970 NFS (micro)[b]	Duncan occupational SEI: $husband_t$ / $husband's\ father_t$		total births, intended births, unintended births	eight[b]	OLS regression
Bean and Wood 1974	1960, 1970 (micro)[c]	own income / income in same race-age-occupation-education group		children ever born, by parity	husband's income, wife's education and age at marriage, rural-urban	OLS regression
Ben-Porath 1975	1968–1970 (micro)[d]	M and F education/ predicted earnings of husband's father		# of children	# of siblings, farm background, indicator for wife's literacy	OLS regression
Behrman and Taubman 1989, 1990	1981/1982 (micro)[e]	income of younger generation/ income of parents		# of children	cousin covariances to control for unobserved preferences (intergenerational serial correlation)	latent variable model
Crimmins-Gardner and Ewer 1978**	1968–1975 (micro)[f]	husband's occupation / wife's father's occupation and spouse's evaluation of family's relative status at t, t-10, t+10		current and desired # of children	husband's income in '69 and '74, wife's potential income, wife's age at marriage	OLS regression

D. Freedman 1963	1955 (micro)[g]	own income / income in same age-occupation-education group	# live births by marriage duration	husband's income, wife's income, her # years in LF, current and future LF status	multiple classification analysis
R. Freedman and Coombs 1966	1962–1963 (micro)[h]	wife's perception of: general adequacy of income; husband's income / income in same age-occupation-education group; and husband's actual income/family income of 3 best friends	# live births between interviews, # expected births	income level, religion, parity	multiple classification analysis
Johnson and Lean 1985	1970 census (micro)[i]	husband's income / income in same age-occupation-education group, by race	# live births to wife	husband's absolute income + ten other[i]	OLS regression
Kunz 1965	1960 census (micro)[j]	husband's income / income in same age-occupation education group	children ever born	woman's age at marriage	mutiple classification analysis, correlations
MacDonald and Rindfuss 1978**	1970 NFS (micro)[k]	wife's response to: better off at marriage than parents when an adolescent?	# planned births in first 3 years of marriage, wanted and unwanted fertility, # live births, desired family size	wife's education, race, religion, # sibs, age at marriage, res. on farm; husband's father's SEI; marriage duration	OLS regression
Maxwell 1991	1945–1985 (micro)[l]	median wages $_t$/ median wages $_{t+20}$ (t=year of resp.'s birth)	age at first birth	female wage, predicted income at age 30, race, period controls	OLS regression
Olneck and Wolfe 1978**	~1948–1970 (micro)[m]	sibling difference in 1973 monthly earnings	sibling difference in # of children	sibling differences in age, age^2, years of education, current marital status, cognitive skill	OLS regression

Table 6 (continued)

Study	Years covered	Relative income measure	Relative cohort size measure	Fertility measure	Other variables	Method of analysis
Oropesa 1985	1972–1983 (micro)[n]	responses to questions re: own and parents' relative affluence		expected family size: next 5 years and total	nine[n]	multiple classification analysis
Reed, Udry and Ruppert 1975	1969–1970 (micro)[o]	husband's ann. income / median family income in respondent's race and city, husband's annual income/mean family income in husband's race-education-occupation group		# of live births	respondent's age at marriage, age at first birth, duration of marriage	multiple-partial correlation analysis
Stevens 1981	1962 (micro)[p]	(S-$\bar{S}$), ($\bar{S}$-F) and ($S + F$) S = son(respondent)'s SEI, $\bar{S}$ = mean SEI in son's origin cohort, F = father's SEI		completed family size	H and W farm origin and education, marriage duration, wife's age, husband's age at marriage	OLS regression
Thornton 1979 [**]	1955–1970 (micro)[q]	husand's income/income in same age-occupation-education group, by region; subjective: current vs past and future, current vs friends and relatives; current vs parents; response to "How well-off do you feel? and "How well off expect to be in future? and as in MacDonald and Rindfuss, 1978		current parity, # expected in future	1) religion, education, marital duration, age at marriage, (parity) 2) ideal family size, contraception attitude, child quality tastes 3) FLFP	multiple classification analysis
Thornton 1979 [**]	1975 (micro)[r]	six[r]			1) marital duration, religion,race,(parity) 2) education, fertility control, value of children, age at marriage	multiple classification analysis

Thornton 1980**	1968–1974 (micro)[s]	huband's income/ parental head's income, parental welfare indicator	cumulative childbearing to '72/ 4, expected and ideal family size	year married, race, religion, H&W education	OLS regression
Westoff and Ryder 1977**	1970 NFS (micro)	husband's income/ average income of others in his education group	total, wanted and unwanted births, women 35–44	none	multiple classification analysis

** indicates a study which provided little or no support for the Easterlin hypothesis

[a] 520 Seattle males married only once, with spouse present.

[b] 2,265 women from 1970 National Fertility Survey, married only once and intending to have no more children. Additional variables were: wife's education, wife's age at marriage, husband's age, farm background, current farm residence, religion, race, length of exposure to risk of pregnancy (in equations for unintended pregnancies).

[c] One percent 1960 and 1970 census data for white, Black and Hispanic women in five southwestern states: once-married, aged 40–49, age at marriage not over 29.

[d] 740 male family heads in the first three waves of the PSID (restricted to those whose wife was white and non-Catholic, and at least 35 years old, to approximate completed family size)

e Over 3,700 respondents in a follow-up survey to the NAS-NRC Twin and Adult Offspring Survey.

[f] 118–152 previously unmarried couples, married in the summer of 1968 in Peoria and Decatur, Illinois, with husband 30 years of age or under at marriage (*N* changes over time by attrition).

[g] National probability sample of 433 non-farm "fecund planners" married at least 5 years. All variables were categorized and represented in the analyses by a series of dummy variables.

[h] Stratified probability sample of 1,113 Detroit-area white married women who had had a first, second, or fourth birth in July 1961, interviewed three times 1962–1963.

[i] One percent PUS: in CA, 18,248 white, Japanese, Chinese and Black women in same-race marriages aged 40+; in Hawaii, 443 white, Japanese and Chinese women in same-race marriages aged 40+. Additional variables were couple's nativity, central city residence, and extended familism; wife's marital stability, age at first marriage, education and LF status; husband's education and three categorical LF/occupational dummies.

[j] 3,977,487 white women 35–44 in the 1960 census 5% sample, married once, husband present, non-farm.

[k] 3,192 once-married, currently married women who were living with both parents at age 14, married to once-married men who were residing with both parents at age 14.

[l] 1,314 respondents from the NLS (both older and younger women) in three birth cohorts, 1923–1929, 1930–1937 and 1944–1954.

[m] 352 weighted pairs of brothers who had attended sixth grade in Kalamazoo, MI between 1928 and 1950.

[n] 500 men and 605 women aged 20–29 in their first marriage, and not married at or before age 16, from the NORC General Social Survey. Other variables used were "perception of past finances" ('origin'), "perception of current finances" ('destination'), age,marriage duration, education, religion, number of siblings, (all for the respondent), and year.

[o] 495 Black and 809 white ever-married women aged 15–44 taken from a random sample in 16 SMSAs. Respondents had to be married once, spouse present, nonmenopausal and nonsterile.

[p] white males aged 35–64, in their first marriage, taken from the "Occupational Changes in a Generation Study", a special supplement to the March 1962 Current Population Survey.

[q] 1955 GAF: 1,805 white, currently married with no a) fecundity impairment, b) pre-marital births or c) unwanted births.
1960 GAF: 1,848 aged 18–39/44 and as above.
1965 NFS: 2,355 under age 55, and as above.
1970 NFS: 3,003 under 45 and as above.

[r] Approximately 1,500 men and women aged 18 and over, with analyses conducted separately for those under and over 40, with restrictions as in *q* above.. For those over 40, only those whose first marriage had remained intact for at least 15 years. The relative income variables used in the analysis are described in the appendix.

[s] 566 couples from the PSID: married after 1968 but before 1971(for whom the dependent variable was cumulative fertility to 1972) and 1973 (cumulative fertility to 1974)

census, and a lack of variance in the income variables across states in 1960 and 1970. Another concern he addressed was that, as a *per capita* measure, his past income variable's significance simply reflected the effect of past fertility on present fertility. He tested this by including in the regressions each state's child-woman ratio twenty years earlier, and found that past income experience still turned out to be highly significant with a coefficient not greatly reduced in magnitude.

Moffitt (1982) tested a 'life cycle' version of the hypothesis, using published 1947–1977 Current Population Survey (CPS) data to construct a predicted 'lifetime wealth' figure for each cohort, and using the ratio of own wealth to that of the cohort twenty-five years older, in equations for age specific fertility rates for women 20–24, 25–29, 30–34, 35–39 and 40–44 (giving him 145 observations). His results "provide strong support for the life-cycle model of the Easterlin hypothesis... However, the Easterlin hypothesis does not provide a total explanation of the baby boom and bust... there appear to be underlying trends in fertility rates that are independent of relative wealth" (p. 250).

Macunovich (1996a, 1998) responded to the idea of "underlying trends" independent of relative income by incorporating the female wage along with relative income in models describing the fertility and labor force participation rates of African-American, white and all women aged 20–24. She aggregated CPS microdata for the period 1963–1995 in order to try to construct variables consistent with Easterlin's theory, and also to develop an 'exogenous' female wage (that is, a wage free of the effects of changing levels of education and experience). Her relative income variable was the average expected earnings of all unenrolled men in their first five years of work experience in time *t–1* relative to the average family income of all families with children under 18, with head (of either sex) aged 45–54, in time *t–6*. She included, as well, an interaction term between male relative income and the female wage, under the assumption that the net effect of the female wage would have been changing over time because of women's increasing levels of LFP (an increasing income effect). Her estimated equations all fit the data extremely well (R^2 of 0.99 and above) – even in differenced form – with little or no serial correlation. The model appeared to fit African-American fertility even more closely than white – and this group's fertility was found to be sensitive not only to African-American parental family income, but also to parental family income in the white community: a 'social mobility' effect. In addition, like Ahlburg (1982), Macunovich (1997a) found an extremely close relationship between her male relative income figure and relative cohort size. But she found both relative cohort size and *position* effects (with cohorts on the trailing edge of the baby boom hit hardest by cohort effects because of declining aggregate demand) – and found that military enlistments and the trade balance had introduced 'noise' in the relationship over time, raising male relative income during the Vietnam war and lowering it in the 1980s.

Shields and Tracy (1986) tested two models for TFR: one for 1920–1980, and the other for 1946–1980. Both used a population ratio as a relative cohort size measure, and included measures of real personal income, female labor force participation, and infant mortality. The shorter period included housing costs as well. All of these variables were significant with the 'expected' signs

in the regression for the longer time series, but in the 1946–1980 period, with house prices included, female labor force participation lost its significance and personal income was only marginally significant – although both retained the 'correct' signs – and house price was not significant. The inclusion of the female labor force participation variable in a macrolevel equation is questionable, however – so to some extent it is gratifying to see the model hold even when FLFP lost its significance. The age structure variable was highly significant throughout, and both regressions explained over 98% of the variance in the TFR. Shields and Tracy concluded that the income and age structure variables combined to account for the dramatic increase in fertility between 1940 and 1960. They felt that, although rising income continued to exert a sizeable upward influence on fertility during the 1960s and 1970s, the age structure variable started to exert a downward influence, which they interpreted as consistent with Easterlin's hypothesis.

Evans (1990) found strong effects of cohort size on both the timing and quantity of cohort fertility, using proportional deviation from trend as his measure of relative cohort size. Like Macunovich (1998), he found that the effect among nonwhites was even stronger than among whites. However, like Shields and Tracy (1986) he included a number of potentially endogeneous additional controls – cohort unemployment rates at age 20–24 and cohort educational attainment and female labor force participation rates – which might have diluted his measured effect of relative cohort size. Even so, he estimated a six to seven month increase in median age at first birth, and a 0.2–0.8 decline in average number of children, for the largest cohorts as a result of their size.

Two other studies – Devaney (1983) and Schapiro (1988) – estimated simultaneous equation systems with aggregate data. Schapiro used a four equation system for age-specific (20–24 and 25–34) fertility, female labor force participation, marriage, and divorce, with age-specific RCS constructed as males 20–24/25–64 and males 25–34/35–64, and age specific RY constructed as ratios of age-specific expected male fulltime earnings. Of four equations estimated for fertility (two age groups, one set of equations using RCS and the other using RY), only RCS for the 20–24 age groups proved to be significant (at the 0.01 level). However, in an otherwise identical system when he added absolute male income (MY) to each equation, he found significant estimates for the 25–34 group when using RY and MY, and for the 20–24 group (and very weakly for the 25–34 group) when using RCS and MY.

Devaney's (1983) analysis responded to Butz and Ward (1979), because she felt that by ignoring relative income as a predictor of female labor supply, Butz and Ward had failed to identify an indirect effect of relative income on fertility. Devaney addressed this shortcoming of their 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.

But Smith (1986) dismissed Devaney's results because of a failure to control for strong period effects in her model: he suggested that the fertility

and income effects Devaney emphasized for younger individuals, in fact were common to the entire population over the period. He felt that there were no strong *period* income effects of an Easterlin or Devaney type (that is, a progressive worsening of incomes at age 20–24 relative to those at later ages) that were not directly related to economic recessions.

Devaney (1986) responded to his criticisms by pointing out aspects of her model formulation which did allow for period effects, and the fact that her relative income variable differed from Easterlin's in that it allowed a cohort's relative income position to vary over the lifecycle, so that his criticism regarding the dependence of its movements on economic recessions was invalid. She concluded with an agreement that more work needed to be done to separate out the effects of cohort and period on fertility in this century, but emphasized that observed period effects themselves might be the *result* of changing population age structure.

6. Microlevel analyses using North American data

Technical details of all of the studies discussed here are presented in Tables 4 (Canadian) and 6 (U.S. microlevel analyses), with asterisks indicating those with the least supportive results.

6.1 Unsupportive microlevel analyses

The often-repeated assertion that the Easterlin hypothesis has not found support in micro level analyses appears to be based largely on the results of seven published studies: Crimmins-Gardner and Ewer (1978: 'CGE'), MacDonald and Rindfuss (1978: 'MR'), Olneck and Wolfe (1978: 'OW'), Thornton (1978: 'T78'; 1979: ';T79'; ;1980: 'T80') and Westoff and Ryder (1977: 'WR'). Three of these seven appeared together, in the first volume of *Research in Population Economics* (1978) – and perhaps for that reason have tended to be given considerable weight in the literature. But a close review of all seven of those studies has left me wondering why they have dominated citations over the years. Perhaps it is because of the wide range of questions – often ingenious – asked of respondents in an attempt to measure their feelings of relative affluence. *(For a list of questions used in these studies, see the Appendix.)* It is, in the end, those questions and their use in the analyses, however, which leave me so uneasy about most of these studies.

Five of the studies – CGE, MR, T78, T79 and WR – adopted very similar methodologies and often made use of very similar survey responses in formulating their relative income measures. They are reviewed together first in this section, followed by a discussion of OW and T80.

All five were microlevel analyses which made use of similar survey questions designed to provide what were termed "objective" and "subjective" assessments of relative economic status. Four were based on sample sizes in excess of 1,500, (CGE, however, was based on a sample of only 118–152), and all used similar measures of relative economic status, in categorical form (e.g., 'better off', 'as good as' or 'less well off'). This

type of analysis can be misleading, however, since it implicitly assumes that young adults' minimum consumption threshold (below which they would forego childbearing) is *identical to*, rather than simply *a function of* parental income. This problem is discussed more fully in Sect. 10.7.

In addition, all used endogenous 'subjective' measures of relative economic status (for example, asking "How well-off do you feel?" and assuming that the hypothesis required those who felt most financially secure to have had the most children to date – or to expect the most in the future, on average). However, all other things equal, a couple who chose to have children early *would* have been less well-off in per capita terms, *because of that decision*, than a couple with no children. Similarly, if a woman indicated that she felt less well-off at the time of the interview, but then indicated that she intended to have more children in the future than women who felt 'well-off', this was interpreted as behavior contrary to the relative economic status hypothesis. But in this case it seems likely that the researcher was observing a timing effect in the context of 'target' family size desires. She had had no children because she had thus far felt financially unable to do so, but because she harbored the standard desire for 2.5 children she planned (i.e., "hoped") to have children in the future, whereas the woman who had already borne her 2.5 children, given her feelings of financial security, planned to have no more.

Not surprisingly, the results of these studies were almost uniformly inconclusive or even negative with regard to the relative income hypothesis. The limited support they provided included CGE's evidence of a 'weak' positive relationship between the wife's assessment of her status, and current and desired fertility; T78's finding of a significant positive correlation between feelings regarding current financial circumstances, and planned fertility, in 1965; and T79's finding of a significant negative relationship between a couple's desire for 'non-home' goods (e.g., vacations) and fertility. But none of these studies formulated or tested continuous relative income variables based on husband's current income relative to parents' past income.

The other two of the seven unsupportive microlevel analyses were Olneck and Wolf (1978) and Thornton (1980). Thornton made use of about 560 couples married during the first few years of the University of Michigan's Panel Study of Income Dynamics (PSID). These couples' fertility was reported in 1972 and 1974, and Thornton calculated a continuous relative income variable: the husband's annual income relative to the parental head's annual income or to the parental welfare indicator (all smoothed somewhat, making use of husband's observed income prior to the observed cumulative fertility, and parental head's income averaged during the years when the 'child' was still at home). The results were very mixed: husband's absolute income was estimated with a negative (and significant) effect in 1974 (but positive and not significant in 1972), while parental income was negative in all but one case and significant in three (and thus supportive of the hypothesis) – but only before controls for husband's and wife's education were added. In ratio form, with respect to total expected family size and additional children expected, relative income was significant with a positive coefficient (i.e., supporting the hypothesis) which increased when educational controls were included.

But many questions are raised by this study – most importantly, whose parent was being reported – his or hers – since the PSID provides only one or the other? Would the results vary if cases with wife's parent were analyzed separately from cases which observed the husband's parent? Why was parental head's income used, rather than parental family's income – and why were these not tested on a per capita basis? Was there any control for husband's status as a student during the period? And what information was used to calculate smoothed income figures for those married and/or having a child in the first year of the PSID – who could not have been observed in the parental home?

Olneck and Wolfe made use of a dataset of 352 pairs of brothers, to see whether differences in sibling earnings were correlated with sibling differences in number of children (on the assumption that parental influence was the same for each sibling). They found no significant effects – but to find a significant effect, the researchers would have needed a sample of brothers married to women with identical preferences and parental living standards (and fecundability), as well as similar effects on tastes from other aspects of their lives (peer groups, etc.) These non-parental differences are assumed to be normally distributed in a larger population, so that their effects are captured in the (zero-mean) error term. This type of averaging effect could not occur in a comparison of two brothers, however, so to some extent it would have been surprising if these results *had* been significant.

6.2 Supporting microlevel evidence from North American studies

I was, frankly, very surprised at the number of microlevel studies whose findings support the relative income hypothesis. There are a total of fifteen of these studies – eight using 'parental' relative income variables (and thus directly relevant to the Easterlin hypothesis), and seven using husband's income relative to a 'predicted' income. The quality of these studies is mixed, so that the credibility of their findings varies. But the overall impression they give is that relative income is significant in fertility decisions – with the strength of the relationship varying directly with the sensitivity of the relative income variable(s) used, to Easterlin's basic hypothesis.

One of the most impressive of the microlevel analyses was conducted by Chaudhury in 1977. This was a remarkably detailed and careful study – especially given its vintage – using microdata on approximately 304 once-married women under age 46 interviewed in Toronto in the 1967/1968 Canadian Family Growth Study. The primary drawback of the study was his use of a relative income measure calculated as husband's actual versus predicted earnings (on the basis of age, age at marriage, education, occupation and place of birth), rather than actual versus parental. It is likely that this was because of data constraints. This measure of relative income was calculated both currently and at the time of marriage, and the effects of both measures were tested on the length of the first and second birth intervals and on total births controlling for duration of marriage, wife's age at marriage, and a number of other factors (see footnote *d* in Table 4).

He found a positive relationship between relative income and fertility among women with no premarital conceptions – a relationship which was more pronounced, as hypothesized, among those married 0–4 and 5–9

years. And yet, consistent with the relative income hypothesis, he found no relationship between relative income and *desired* family size, with low, medium and high relative income groups desiring 3.01, 3.01 and 2.99 children, respectively.

In his sample he found that patterns of contraceptive use and FLFP were highly consistent with Easterlin: all zero parity couples in the low income group were using contraception, as compared with only 54.5% of high relative income couples, and the FLFP rate among women in the low income group was more than 50% higher than in the high relative income group. All in all, a remarkably complete and consistent set of findings in support of the relative income hypothesis. Its drawbacks, like those of so many analyses, center on the use of categorical variables (a procedure necessitated by multiple classification analysis), and the absence of an intergenerational relative income measure.

6.2.1 Supportive microlevel studies using non-parental relative income measures. Deborah Freedman (1963) appears to have produced the first (published) formal multivariate relative income analysis, in which she interpreted income as relative to the predicted mean among those sharing the husband's age, education and occupation levels (using Duncan's SEI). She used a probability sample of 433 non-farm women, and converted all variables into categories which were represented in the analyses by a series of dummy variables. She found a significant and positive effect of husband's absolute income with a significant negative effect of his relative income: the lowest relative income group had 0.50 children less than the average while the highest had 0.28 children more than the average.

Perhaps the weakest of these analyses was Kunz (1965), undertaken in response to Freedman's analysis, using a subsample of over three million white, married, nonfarm women from the 1960 census five percent sample. He prepared data on husband's income and number of children ever born classified by wife's age at marriage and husband's occupation/education group, and then examined simple correlations within cells. He reported nothing more for each cell than "positive," "negative" or "blanks indicate no relation", and found a total of 18 out of 19 cells showing a 'positive relationship' for wives married at age 22 and above. Most problematic here is the lack of any analysis of statistical significance, or indication of the criterion used to qualify as a "positive correlation".

Bean and Wood (1974) then used census PUS data for five southwestern states – for 1960 and 1970 – to examine relative income effects within races on completed fertility. Their RY measure was once again husband's income relative to that expected for men of his same race-age-education-occupation group. They found fairly mixed results, possibly because of the small number of control variables used, but in general there was a positive effect of RY for whites and Hispanics (but significant only at parity three and four, and in some cases only barely significant) while for blacks the only significant effect was a positive one at parity one in 1970. Actual income, when significant, generally had a negative sign, which the authors attributed to a high degree of collinearity between the two income measures.

More definitive results were found by Reed et al. (1975), who used survey micro data for Blacks and whites from 16 different SMSAs in 1969–

1970. They analyzed the relationship between number of live births and husband's relative income when the latter is measured in two different ways: husband's annual income/median family income of families of the same race in the same city, and husband's annual income relative to the mean family income of families whose head is in the same race-education-occupation group as the respondent's husband. The two measures were highly correlated, and thus performed very similarly in all of the analyses. Controlling for respondent's age at first birth and duration of marriage, they were able to explain about 45% of the variance in fertility for each race. The differences in number of births between those in the top and bottom quintiles of relative income, controlling for the same variables, was about 0.8 for Blacks and 0.5 for whites.

And finally in this group, a somewhat back-handed tribute to the relative income hypothesis was paid by Johnson and Lean (1985), in testing a social mobility hypothesis which expected *high* male relative income in minority groups to *reduce* a couple's fertility. Using nearly 19,000 women from the 1970 census microdata for California and Hawaii, they developed relative income figures using a husband's own income relative to the predicted average wage in his race-age-education-occupation group. Their results resoundingly supported the (positive) relative income hypothesis for whites in all cases in California and Hawaii and for Japanese and Chinese in Hawaii. There was no significant relationship for the minority groups in California.

6.2.2 Supportive microlevel studies using parental relative income measures. A few of these studies suffer from a problem mentioned in Sect. 6.1 and discussed in Sect. 10.7, namely, the use of categorical measures which implicitly assume that second generation minimum consumption levels are *equal to*, rather than *a function of* parental income. But because any bias introduced would tend to be a negative one, such shortcomings don't negate positive findings – they just indicate that the findings might have been stronger if measured without the bias.

R.Freedman and Coombs (1966) conducted an analysis similar in many ways to those reviewed in Sect. 6.1, with a sample of 1,113 women, and 'subjective' measures of relative income: whether the woman considered the family's income to be "enough to get along on comfortably", adequate with respect to others of the (husband's) same age, education and occupation, and whether she considered it adequate relative to that of her (previously named) three best friends. They concluded that perceived adequacy of income had a consistent effect on both expected family size and speed in reaching a given stage in the family life cycle. However, no definition of "adequate" was provided.

Bahr, Chadwick and Stauss (1975) developed five different relative status measures for 520 Seattle men in their first marriages with wife present. They asked respondents to rank

1) their own current lifestyle relative to that of
   a) their parents when they were the same age
   b) others with their own level of educational attainment; and

2) their own occupational status relative to
   a) that of their fathers
   b) their own desired level.

These four 'relative status' measures, along with a ratio of current family income to the average income of others of the same age, education and occupational status – a 'relative income' ratio – were examined in terms of their effect on number of children, controlling for age, marriage duration and number of siblings. They found strong and significant positive effects of their relative income variable in all cases except those (84) males with four or more children, and they concluded that relative income is a more important determinant of fertility than absolute income.

In the same year, Ben-Porath (1975) used male family heads in the first three waves of the PSID (restricted to those whose wife was at least 35 years old, to approximate completed family size) to explore relative income effects on total number of children. With 'first generation' information only on the man's father (i.e., none on the wife's parents) he predicted the father's earnings using his years of education, occupation, age and age squared, and indicators for rural-urban and south-non south.

Unlike Thornton (1980), (in Sect. 6.1), who did not differentiate between husbands' and wives' fathers, Ben-Porath found a significant (negative) effect of first generation earnings on the number of children in the third generation – despite the consistent insignificance of second generation earnings, wages, or non-labor income – and a positive effect of number of siblings in the second generation. However, in estimating his 'full' model (including controls for both generations) he used only education of the husband and wife in the second generation as the presumed 'numerator' of a relative income variable. He concluded that the number of children in the third generation was inversely related to the predicted income of the first generation and directly related to the number of siblings in the second generation.

A novel twist was then placed on Easterlin's hypothesis by Bean and Swicegood (1979), who disaggregated the births among 2,265 respondents to the 1970 NFS survey into two categories – intended and unintended – which they analyzed along with total births. They refer to their relative income measure in terms of intergenerational mobility, and constructed it as the ratio of the Duncan SEI score of the woman's husband's occupation to that of her husband's father. They found strongly significant effects, but only on *unintended* births:

> "the upwardly mobile reveal a level of unintended childbearing that is nearly 0.2 of a child above that of the nonmobile couples, and the downwardly mobile a level nearly 0.3 of a child below . . . [T]hese results lend support to the relative economic status hypothesis . . . Interestingly, however, the operative mechanism does not appear to be one associated with the demand for births. *Rather, it would seem to be one associated with the relaxation of the need for fertility regulation*" (pp. 616–617, emphasis added).

Stevens (1981) used microdata on the occupational SEIs of fathers and sons in an attempt to decompose social mobility into two components: the

son's occupational SEI score *(S)* relative to $\bar{S}$ – defined as the average SEI of sons from his socio-economic origins – and $\bar{S}$ relative to the father's SEI score *(F)*. Thus it was intended that

$$(S - F) = (S - \bar{S}) + (\bar{S} - F). \tag{1}$$

She termed (S–$\bar{S}$) "relative mobility" and ($\bar{S}$–F) "mean mobility", and expected relative mobility to exert a negative effect, and mean mobility a positive effect, on completed family size. In order to control for the origin and destination *levels*, as well as differences, she wanted to include four variables in her regression: *S, F, (S–*$\bar{S}$), and *(*$\bar{S}$–F), and to avoid an underidentified equation, she constrained the coefficients on *S* and *F* to be equal, entering them as the variable *(S+F)*. Her results were all highly significant, but with the 'wrong' signs, and she was able to explain only a portion of the variance – on the order of 2–3%.

It would seem, however, that in order for her equation (1) to hold, she would need to constrain the coefficients on the right hand side variables to be equal. As estimated they are of opposite signs, with the coefficient on the second RHS variable negative, and its absolute value approximately twice that of the other coefficient. Thus, instead of estimating the effect of *(S–F)*, she estimated the effect of *(S– F/*2–$\bar{S}$/3) – and she did not report any test of the validity of her constraint on the levels of *S* and *F*, so that it is difficult to interpret her strongly significant results.

Oropesa's (1985) analysis suffers generally from the same problems as MacDonald and Rindfuss (1978) – but here the situation is further complicated by the use of fertility *expectations* (short term and total) as the dependent variable. There is a notorious lack of consistency between intentions and actual fertility. But by far more difficult is the use of a 'relative affluence' measure constructed using the answers when respondents were asked, with regard to their parental family when they were 16, and then with regard to their own family currently, "Compared with American families in general, would you say your family income is far below average, below average, average, above average, far above average?" Oropesa constructed a 25 cell 'origin-destination' matrix, which was then collapsed to seven categories: 0) major drop, 1) minor drop, 2) stable low, 3) stable average, 4) stable high, 5) minor rise, 6) major rise.

So once again we have some confusion about the relationship between second generation aspirations and first generation income, muddied still further by the use of categorical variables. Because Oropesa used a measure with a broader range of categories, the results here are somewhat stronger than in the MacDonald and Rindfuss study, but only when Oropesa focused on those in categories (0), (1), (5) and (6) – the extremes of his range.

Behrman and Taubman (1989, 1990) found a source of intergenerational data in the NAS-NRC Twin and Adult Offspring Sample, supplemented by an offspring survey they carried out themselves in the early 1970s. These datasets are limited, in that the parental 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 ac-

tual number of children was used in Behrman and Taubman's study. Use of expectations is a limitation, given Blake's (1967) demonstration that desired family size remains farily constant over time, while actual fertility fluctuates in response to economic conditions. 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 with another indicating number of siblings: Their model is:

$$n_t = b + \eta I_t - \pi I_{t-1} + \pi n_{t-1} + x_t \tag{2}$$

where n_t is the completed family size of generation t, I_t is the income in generation t, and the coefficients π on I_{t-1} 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.

They 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).

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 for each generation, rather than male earnings only for the younger 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.

Maxwell (1991) conducted an analysis of age at first birth, making use of data on 1,314 women from the NLS. She was particularly interested in identifying racial differences and determining the relative strengths of aggregate versus individual characteristics on women in three different birth cohorts (1923–1929, 1930–1937 and 1944–1954). Her measure of relative income was the aggregate wage in the year of the woman's birth over the aggregate wage twenty years later, and thus put the 'father's' wage in the numerator (so that a negative coefficient would support the hypothesis). She found that this version of Easterlin's relative wage worked as expected for the 1923–1929 and 1944–1954 cohorts. Indeed, the standardized coefficient on her relative wage variable for each of those two cohorts was four to eight times as large as those on individual level variables, with *t*-statistics of 17 for the earlier cohort and 39 for the baby boom cohort.

7. The theory of population cycles

Meanwhile, as all of these empirical analyses have searched for evidence of the Easterlin hypothesis, a related strand of the literature has attempted to determine, as Kenneth Wachter (1991) put it: "Are there demographic feedback models capable of producing the appearance of cyclicality seen in series of twentieth century U.S. and other first-world births?" (p. 109). Are such cycles theoretically possible, and do the U.S. cycles conform to an acceptable theoretical model? Wachter pointed out that such theoretical models are needed to assist in identifying appropriate age groups and time lags, which in turn would help measure the role of imperfect substitutability in demoeconomic feedback loops. These models would also help in sorting out the cohort vs. period controversy, and bridging the perceived gap between micro- and macro-level findings regarding the hypothesis.

Lee (1974), defined a family of feedback models which might encompass an Easterlinian feedback mechanism, identifying two types of cycle which could be generated – short-term or transient cycles, and longer-term sustained cycles. In that paper he concluded that parameter values for what he termed a 'cohort' model estimated from the United States experience between 1917 and 1982 could not sustain longer-term cycles. However, he felt that such longer term cycles could be generated by what he termed a 'period' model, in which period fertility depends on total labor force size. Frauenthal and Swick (1983) then identified limit cycle oscillations in twentieth century U.S. fertility patterns by extending Lee's cohort model into a continuous-age version. But Wachter and Lee (1989) found detrending errors in Frauenthal and Swick's work, and in re-estimating Lee's 1974 model found that it led to "cycles of implausible period". At this point, while transient cycles looked possible, sustained cycles seemed ruled out.

In parallel with this work, Samuelson in 1976 developed what he termed "an oversimplified version of the Easterlin theory" using a two-generation overlapping generations model in order to subject it to "rigorous analysis" – a model which Feichtinger and Sorger (1989) extended to a continuous-time model. Using nonlinear differential equations they were able to generate an Easterlin cycle with a period of about 43 years. They

pointed out that a discrete-time framework was more appropriate for describing population dynamics but "the discrete time approach does not yield the period length of the Easterlin cycle."

Denton and Spencer (1975) and Anderson (1982) presented simulation models examining the cyclical implications of an Easterlinian model, and Day et al. (1989) and Feichtinger and Dockner (1990) attempted to incorporate economic variables into theoretical models, demonstrating in the first case that "nonperiodic, essentially unpredictable demoeconomic behavior is robust" and in the second that swings in economic and demographic variables will occur, largely as a result of intertemporal substitution between current and future consumption. Neither of these latter two models incorporated age-structure effects, however.

And most recently, Chu and Lu (1995) again took up the models specified by Lee (1974), testing an unrestricted version which incorporated both Lee's 'period' and 'cohort' models, and found a limit cycle solution, but it was not stable so that the population trajectory would not converge to that limit cycle.

Wachter (1991) attempted to determine the characteristics of a "viable feedback model" and concluded that, "there *are* viable feedback models for U.S. births, but very few, and they are very special" (p. 124). He concluded that

> "the best of the relative size models, like the best of the potential labour force segment models and cascade models [incorporating 'bandwagon' effects in which 'pacers' affect the timing of fertility in a number of age groups simultaneously], come close enough to matching the targets that fine-tuning within the constraints already identified holds out promise of success" (p. 125).

8. The formation of aspirations

Easterlin has long emphasized the need for research into the formation of material aspirations, as a necessary input to studies on the relative income hypothesis. He has explored the concept himself in his classic article "Does Money Buy Happiness?" (1973b), and in two more recent applications of the same concept (1995b, 1996). My attempt here will be simply to give a taste of the literature and the individuals who have been active in this area over the past three decades. It goes without saying that three of the earliest forerunners in the field are Duesenberry (1949), Modigliani (1949) and Leibenstein (1950). Brown (1952) then demonstrated 'habit persistence' econometrically, based on Duesenberry and Modigliani but using Canadian data to show that consumers' reaction to changes in income takes place gradually. "[T]he habit persistence effect is produced by the *highest previous level of consumption experienced.* This effect proceeds undiminished until a higher level produces a new standard." From there, it seems that two nearly separate strands of research emerged, one on either side of the Atlantic.

In the United States, Robert Pollak and Terence Wales have been the dominant figures in the literature on preference formation over the last three decades. In an early work (1969) they estimated a linear expenditure

function based on the Klein-Rubin model (1947–1948) and first estimated by Stone (1954), which allows for a 'necessary' level of consumption before allocating more discretionary expenditures. Their model was estimated for pre- and post-World War II periods, testing various assumptions about preference formation. They found that in the postwar period two different specifications of preference formation were significant – one based on a time trend and the other on 'habit formation' in which current consumption is a function of consumption in the previous period. Wales (1971) similarly found a 'habit formation' model appropriate using Canadian durable goods expenditure data for the period 1947–1968. Pollak (1970) explored some of the constraints which must be placed on a dynamic model with interdependent preferences, if it is to be mathematically tractable.

Pollak's two 1976 articles dealt with theoretical issues in the specification of individual preference functions whose parameters depend on the consumption of others. In (1976b) he conceded that only a fairly limited class of short-run utility functions yielded long-run demand functions which can be rationalized by utility functions. Then in 1979 Howe, Pollak and Wales tested two alternative assumptions regarding taste formation using a quadratic expenditure system: constant tastes and linear habit formation. Their results in the latter case were inconclusive, however.

Correspondingly, in Europe Bernard van Praag and Arie Kapteyn have dominated the literature on individual welfare functions of income, examining and developing the hypothesis that individuals have welfare functions which depend on a comparison with other members of their social group – and then developing methods of quantifying these individual welfare functions. The individual welfare function concept was introduced theoretically by van Praag (1968), and then tested empirically in van Praag and Kapteyn (1973), where they established a strong element of interdependence in individuals' welfare functions.

Kapteyn et al. (1980) carried this further, assuming that one's preferences on consumption were a function both of contacts with others, and also of habit formation (own previous consumption). Theirs was the first attempt to measure both: they found that the former accounted for 2/3 and the latter for 1/3 of individual preferences. They maintained that "by now there appears to be substantial empirical evidence on the endogeneity of preferences."

van de Stadt et al. (1985) estimated a model containing unobservables using Dutch panel data, and found that their results were consistent with the hypothesis that preferences are entirely relative – although they couldn't rule out the possibility that they were part relative and part absolute. And in 1988 van Praag, Dubnoff and van der Sar presented "a method by which individual judgements of the quantities associated with normative descriptors can be used analytically." This was a method of translating individual standards into general 'objective' standards for use in quantitative analysis which they termed the EQA – evaluation question approach. The technique was elaborated further in van Praag (1991) and van der Sar and van Praag (1993), and summarized in non-mathematical terms in van Praag (1993), and then defended in van Praag and Kapteyn (1994).

Kapteyn (1994) argued against economists' aversion to the use of the 'direct method' of establishing individuals' preference functions (i.e., sur-

veys). He attempted to demonstrate its validity by comparing the results of estimating traditional household cost functions using both direct and indirect (revealed preference) methods to see if they appeared to be measuring the same things. Unfortunately, his results were mixed, with the outcome depending on which set of instruments was used to test the null hypothesis (of no difference between the two methods).

Back on the United States side of the Atlantic, Hayakawa and Venieris (1977) explored the implications of interdependent consumer preferences and determined that while the resulting preference map exhibits smooth indifference curves that are convex to the origin, the marginal rate of substitution is positive and diminishing only within a defined range so that "more does not necessarily imply at least as good as". Ahlburg (1984) attempted to develop a formal mathematical representation of Easterlin's hypothesis, addressing in the process issues such as whether an individual's preferences depended on his own or his parents' lifestyle in the parental home. Frank (1985) produced an entertaining and highly readable and informative compendium of findings regarding the causes and effects – from neurological to political – of our "quest for status". And MacDonald and Douthitt (1992) used a micro dataset to test three different theories of consumption behavior – the Modigliani life cycle hypothesis, Duesenberry's relative income hypothesis and Kyrk's (1953) resource deficit hypothesis. They found that the relative income variables were highly significant, and concluded that the RIH model was the simplest and most powerful of the three in explaining consumer utility .

And finally, mention should be made of Leibenstein (1975, 1976), Lestaeghe and Surkyn (1988), and the conversation between Sen (1993) and Bliss (1993), which are more philosophical treatments regarding the formation and role of aspirations in fertility behavior.

9. Age-period-cohort analyses

Many feel that one of the most compelling arguments made against the relative income explanation of post WWII fertility booms and busts, is the fact that fertility rates of all age groups moved in tandem during this period. For example, Rindfuss et al. (1988) presented regression results which appeared to explain over 90% of the variation in the first-birth probabilities of women aged 25, 30 and 35, using macroeconomic period measures, with no cohort controls. Similar analyses were conducted by Smith (1981); Kelly and Cutright (1984); Wright and Maxim (1987, using Canadian data); and age-period-cohort analyses such as those by Pullum (1980); Brass (1974); Page (1977); Namboodiri (1981); Foster (1990), invariably found cohort and period effects swamped by those of age, and period effects much stronger than cohort.

The most straightforward response to such criticism, however, is that made by Devaney (1986) when she observed that "an interesting implication of Easterlin's hypothesis is that these [period] effects were themselves the result of the relative sizes of various cohorts in the labor market at those times." This implication has been explored in studies such as

Mankiw and Weil (1989), Fair and Dominguez (1991), and Macunovich (1997a, b), which suggest measurable aggregate *demand* as well as supply effects of the age structure of the population. Consumption, housing demand, unemployment, inflation and interest rates are among the factors found responsive to cohort size effects in these studies.

Thus cohort effects have continued to attract research attention, since the pioneering work of Hajnal (1947, 1950a, b, 1959) and Norman Ryder (1965, 1980, 1983), as Ní Bhrolcháin (1992) has complained:

> "[a]lthough Hobraft, Menken and Preston (1982) discuss most of the statistical investigations of fertility considered here and note that they are largely supportive of the period approach to fertility, they do not come to a clearly stated conclusion in its favor."

The Hobcraft et al reluctance to "come down to a clearly stated conclusion in its favor" was based on their distinction between simple additive linear models, like those used in studies cited by Ní Bhrolcháin (as well as Gilks 1979, and Pullum 1980) and more complex behavioral models which introduce nonlinearities and additional variables drawn from theories regarding underlying patterns of behavior. A detailed examination of methodologies and findings led them to question the simpler models:

> "Both Page [1977] and Gilks [1979] noted considerable patterning of their residuals in each period table. . .Thus, despite methodological and substantive disagreements, both of these investigators suggest that fertility corresponds not only to age and period effects, but to marriage duration, marriage cohort, and age at marriage. Their work indicates that models considerably more complex than those specified thus far may be required for adequate analysis of fertility" (p. 25).

> "The conventional [additive statistical] approach is much less suitable for the analysis of fertility [than mortality], except perhaps in natural fertility populations where physiological mechanisms play a dominant role. Once goal-directed behavior is introduced, it is important to base any empirical examination on theories or assumptions about how such goals are formulated and pursued" (p. 31).

Wright and Maxim (1987) approached the testing of the Easterlin hypothesis in Canada *via* a traditional age-period-cohort disaggregation. They regressed the fertility rates of seven five year age groups from 1921–1985 on age and period, and then regressed the residuals of this equation on cohort. They found that age alone accounted for 84% of the variance, and "period effects alone accounted for 60% of the residual in fertility rates after age has been controlled. Likewise, cohort effects account for 45%. It is important to note that after controlling both age and period, cohort effects explain an additional 30% of the residual variance." However, when they plotted the curve of estimated cohort fertility effects they found a *U*-shape which indicated the opposite of the Easterlin hypothesis: that fertility varies positively rather than inversely with cohort size.

This result is not so surprising when the non-orthogonality of period and cohort is taken into consideration. Because of it, the time-pattern of

period and cohort coefficients is highly dependent on the order in which the analysts choose to run their regressions. That is, if the authors had first regressed the fertility rates on age and cohort, and then regressed the residuals on period, their results might have been an equally non-intuitive time pattern of period effects. Namboodiri (1981), for example, alludes to this problem in emphasizing that "one can get an infinite number of solutions for the regression parameters (elements of β). It is not widely known, however, that all particular solutions of β lead to the same (unique) estimate of the sum of squares attributable to the model." Thus, while Wright and Maxim's R^2s and analysis of variance will be correct, their graphs of resulting β coefficients must be suspect.

Another reason for their non-intuitive results could be selection bias inherent in age-period-cohort analysis: no matter how the analysts select their dataset, one of these elements will be truncated. That is, with a rectangular data matrix of age by year, cohorts at the beginning and end will be truncated (that is, not observed at all ages), while a rectangular data matrix of cohort by year will necessarily truncate observations of age groups in some periods. Murphy (1992:149–150) discusses this problem of selectivity bias in his analysis of third births in Sweden.

In addition, traditional age-period-cohort analyses assume that cohort effects are uniform over the entire life of a cohort. For example, age specific fertility must be lower than average at all points in a cohort's life cycle in order for that characteristic to qualify as a cohort effect in such analyses. But it is not clear that the cohort size effects discussed by Easterlin will follow this pattern: large cohorts which suffer fertility-suppressing adverse economic conditions at young ages might display compensatory higher-than-average fertility at older ages, if they encounter favorable period effects at that time. Analyses presented in Macunovich (1998) suggest that this has been the case historically: using standardized (zero mean, unit variance) age-specific fertility rates this analysis identifies marked patterns of "catch-up" fertility behavior which explain the uniform movement of age-specific rates during the United States baby boom. Thus the consistency of cohort effects over the life cycle is an empirical question which requires further research – research which cannot be conducted using traditional age- period-cohort analyses.

10. Conclusions

This article has attempted a comprehensive review and critique of the literature regarding fertility aspects of the Easterlin hypothesis. Some one hundred eighty-five published articles and books have been reviewed, including seventy-six empirical analyses of the hypothesis and closely-related concepts (summarized in Tables 1–6), and – although results are certainly mixed – it would appear that an inordinate amount of weight has been placed on a relatively small number of studies with unfavorable findings. In the North American context, there have been an equal number of micro- and macro-level analyses (22), and two-thirds of those in each group (15) have produced favorable results.

What might account for the variation in findings among these analyses? One most obvious reason is that, because of data limitations and idiosyncratic interpretations of the hypothesis by individual researchers, many of the studies with unfavorable findings have been only peripherally related to the Easterlin hypothesis. The following discussion attempts to highlight those differences among studies, which appear to be related to differences in their findings on aspects of the Easterlin hypothesis. Listed below are characteristics which seem to be more typical of unsupportive than supportive studies – although it must be emphasized that often this is a matter of degree. In *aggregate analyses*, studies which find little or no support for the Easterlin hypothesis tend to be those which:

– Use variables which are not *age-specific*. (This is particularly true of the European analyses: *none of the aggregate studies of countries outside North America have used age-specific relative income measures).*
– Use relative cohort size or relative income variables without any other controls.
– Attempt to fit the Easterlin model with older (age 30+) age groups.
– Treat family income and male earnings as interchangeable.
– Use relative cohort size rather than relative income as the independent variable (especially in later years).

In *microlevel analyses*, studies which find little or no support for the Easterlin hypothesis tend to be those which:

– Treat family income and male earnings as interchangeable.
– Use only the husband's characteristics in formulating relative income variables, without information on his or his wife's parents.
– Focus the analysis only on women in intact first marriages with no 'unwanted' or 'unintended' births, and analyze fertility with age at marriage held constant.
– With categorical rather than continuous measures of relative income, treat the second generation's minimum consumption threshold (i.e., the level of affluence required before they feel able to support a family) as *equal to* first generation income, rather than a *function of* it.
– Use expected or desired rather than actual fertility.

Each of these characteristics is discussed in more detail in sections 10.1 – 10.7.

10.1 Age-specific variables

A central mechanism in the Easterlin hypothesis is imperfect substitutability between *older* and *younger* workers, which leads to a widening of the gap between the earning potential of *older* and *younger* workers – *at a given point in time*, or at points very close in time, rather than at a given age. This gap between older and younger workers is used in aggregate analyses as a proxy for the widening gap between the earning potential of fathers and their sons, and observation at points close in time is needed to contrast the lifestyle available very recently to young men and women in their parents' homes, and the lifestyle the young man can achieve on his own.

The summaries in Tables 1–6 indicate very frequent use in aggregate analyses of *two different lags of the same time series*, in place of age-specific relative income measures: e.g., the average expected wage of all males, the general unemployment rate, or per capita GNP. These measures tell us little about intergenerational effects at a given point in time – rather, they tell us how the population as a whole at any point in time rates its earning potential relative to that of the population as a whole at an earlier point in time. This *is* a measure of 'relative income', but not Easterlin's preferred measure, and not a measure one would necessarily expect to affect changes in birth rates. These non-age-specific measures tend to be more successful in North American than in European analyses – an empirical result which merits further research.

In European analyses, Ermisch (1988a,b) came close to an age-specific variable in using a cohort-specific measure (a ratio of own cohort size to the size of the largest – 1948 – cohort). It is noteworthy that he appears to have estimated the strongest and most significant cohort size effects in the studies using this cohort-specific measure. But this measure would have captured the full labor market effects of relative cohort size only in the unlikely case that the size of older cohorts was unchanging, or changing in the same pattern as that of younger cohorts. Similarly, relative cohort size variables based on a ratio of younger to older (or vice versa) age groups have always been found significant in North American analyses – at least in the period up to about 1980 – and in the majority of analyses for countries outside of North America.

In the North American context, the measures which have come the closest to Easterlin's preferred concept are Moffitt's (1982) ratios of cohort wealth, Easterlin's own ratio of the income in young families relative to the income in families of their parents' generation (used in his own later analyses and also by Abeysinghe 1993; Butz and Ward 1977; Lindert 1978; Schapiro 1988), and Macunovich's (1996a, 1998) ratio of young men's expected earnings to older families' income. All of these analyses produced strongly supportive results. On the other hand, results might be sensitive to the formulation of the age-specific variable, as suggested by the conflicting results of two U.S. studies conducted at the state level using census data (DeFronzo 1976; Lindert 1978). Kenneth Wachter (1991) emphasized that theoretical work exploring the possibility of sustained population cycles is needed in part to assist in the identification of 'correct' age groupings.

10.2 Controlling for other factors

It often seems that there is a strong 'all or nothing' mentality operating with regard to relative cohort size/income effects. This seems to be the case particularly in studies outside of North America, where researchers who find that a crude test (e.g., of relative cohort size vs. TFR) has not 'worked' – or that it appears that other effects are operating as well – then fail to test other formulations to quantify the several effects, or move on to other formulations which omit relative cohort size or income altogether.

Easterlin emphasized that the underlying mechanism in his hypothesis – the relationship between relative cohort size and relative income –could be expected to operate only in a relatively closed economic-demographic sys-

tem. Immigration would weaken the relationship by supplementing small cohorts when there was an excess demand for labor and thus diminishing the beneficial effects of small cohort size on relative wages. As a result, tests of the hypothesis must control for any 'leakage' in the system, whether it occurs in the form of immigration or international trade.

Similarly, the effects on fertility of female wages and labor force participation must be considered along with male relative income, as demonstrated by Macunovich (1996a, 1998) – but at the same time, allowance must be made for the potential endogeneity of these variables: the effects of male relative income on female labor force participation, and in turn on the observed female wage. And, as demonstrated by Pampel (1993), in cross-country analyses it is imperative that allowance be made for institutional differences which might ameliorate any effects of relative cohort size on relative wages, or relative wages on fertility.

10.3 The Easterlin model and older age groups

Easterlin's emphasis was on young couples in the early stages of household formation. He did not suggest that parents are the *sole* influence on their children, but rather that they will exert a significant and measurable influence, primarily at younger ages – presumably with other influences assuming dominance later in life. A mechanical application of his model to older age groups, however, very unrealistically implies that parents are the only influence, and that this effect should not weaken over time as individuals move further from their origins. It also makes no allowance for the possibility of 'catch-up' behavior among cohorts who are economically disadvantaged early in adulthood.

10.4 Male earnings vs. total family income – in both aggregate and microlevel analyses

Valerie Oppenheimer (1974, 1976) has emphasized the importance of using family income in the denominator of a relative income ratio, to describe the parental standard of living, but male earnings in the numerator to understand the effects of severe deterioration in young men's earning potential. In the Easterlin hypothesis, when young men's relative earnings fall demographic adjustments will include delayed or foregone marriage, and increased labor force participation among wives when marriage occurs. Family income in the numerator of a relative income variable is endogenous, and its use will severely hamper our efforts to measure the effects of declining relative income. Similarly, children of baby boomers tend to come from two-earner parental homes, so that the use of only the father's earnings in the denominator of male relative income may understate the material affluence of an individual's home environment.

Failure to differentiate between husband's earnings and family income is a particularly acute problem in microlevel analyses using subjective measures of well-being (e.g., Crimmins-Gardner and Ewer 1978; MacDonald and Rindfuss 1978; Thornton 1978, 1979; Westoff and Ryder 1977). It is possible that respondents answer that they feel (or felt, early in their marriage) well-off *because they have chosen dual labor market participation* in

order to improve their financial position. Use of such endogenous measures would have produced lower mean fertility among 'well-off' couples than among those not so well-off (but with a stay-at-home wife).

10.5 Characteristics of the husband or the wife?

Problems may occur in microlevel analyses when, for example, it is assumed that brothers with the same parental background should have similar fertility, controlling for absolute income (as, for example, in Olneck and Wolfe 1978). However, to find a significant effect, the researchers would need a sample of brothers married to women with identical preferences and parental living standards (and fecundability), as well as similar effects on tastes from other aspects of their lives (peer groups, etc.)

Another source of problems occurs in the use of datasets like the University of Michigan's Panel Study of Income Dynamics (PSID), which contain information on either the husband's or the wife's parents, but not both. Thornton (1980) failed to differentiate between these two types of cases. His inconclusive results should be contrasted with the positive PSID findings of Ben-Porath (1975), where tighter controls appear to have been used in the analysis.

10.6 Marital status and 'unwanted' or 'unintended' births

The omission of women with 'unintended' pregnancies may censor observations at higher levels of relative income if – as suggested by the findings of Bean and Swicegood (1979), "the operative [relative income] mechanism does not appear to be one associated with the demand for births. Rather, it would seem to be one associated with the relaxation of the need for fertility regulation" (pp. 616–617). That is, a more economically favored individual may feel less motivation to contracept effectively. To her, the cost of another child is not high enough relative to the (psychic as well as economic) cost of strict contraception.

At the other end of the spectrum, however, the tendency to restrict analyses to once-married intact couples may omit from the analysis many of those most economically disadvantaged: we may be observing only (or disproportionately) those who felt sufficiently well-off to marry and start a family – and who experienced less divorce/separation induced by financial stress. And controlling for age at marriage eliminates any effect of the relative income hypothesis operating through delayed marriage. Any of these omissions would tend to weaken the measured relative income effect by focusing the analysis on the most stable segment of the population, where fertility rates fluctuate least. One of the very useful aspects of the Easterlin hypothesis is that it posits a relationship between marital and out-of-wedlock births: reduced male relative income tends to reduce overall fertility, but to leave a larger proportion of young women exposed to the risk of out-of-wedlock births by reducing marriage rates.

10.7 Second generation's minimum consumption threshold: equal to first generation income?

A common characteristic of five of the seven microlevel analyses frequently cited as evidence against the Easterlin hypothesis – Crimmins-Gardner and Ewer (1978); MacDonald and Rindfuss (1978); Thornton (1978, 1979); Westoff and Ryder (1977) – appears to be the assumption that second generation minimum consumption thresholds are *equal to* first generation income. However, as many critics of the hypothesis have pointed out (e.g., Smith 1981), most young people do not expect to earn in the early years of marriage what their parents had been earning only a few years before. This is an obvious point, and one which is entirely consistent with the hypothesis. It is more likely that young people's minimum consumption thresholds will be a *function of* parental income – with a distribution of young people around some mean probably located below the mean parental income. Thus knowing whether or not a young couple is currently earning more or less than parents a few years earlier tells us *very little* about *whether they consider that current income adequate to support a family.* But in nearly all cases the relative status variables used by these researchers were categorical: often with only three categories – 'better off', 'same', and 'less well-off' – so that if a woman indicated in her response to any question that she felt 'less well-off' (than parents, friends, other relatives, or herself at an earlier date) it was assumed that her fertility should be lower than a woman who felt 'well-off'.

In such a restricted categorization there would be two types of 'deviants' obscuring any relationship between relative income and fertility: those whose income is less than their parents' but greater than their minimum consumption threshold (who would tend to have children even though the researchers' categorization implies that they 'shouldn't'), and those whose current income is greater than their parents' but less than their minimum consumption threshold (who would tend not to have children even though the researchers assumed that they 'should').

10.8 Additional research: is it justified?

The weight of supportive analyses certainly seems to justify further work in this area. It would seem that simplistic interpretations of the hypothesis early on led to unrealistic expectations regarding the potential replication of population cycles in the United States – and when the 1980s failed to echo the 1950s the tendency was to abandon the hypothesis altogether. These simplistic interpretations either ignored the mechanism of transmission suggested by Easterlin, and examined only the link between relative cohort size and fertility, rather than relative income and fertility, or they made no allowance for the effect of factors other than cohort size on relative income and hence on fertility.

It might be advisable in future research to distinguish carefully among at least four different components of the hypothesis as it relates to fertility:

– the relationship between relative cohort size and relative wages: is there a measurable effect, and if so, is it only a supply effect, or is there a

demand component – and how important is it relative to other factors such as technological change and international trade?
– the relationship between relative income and the fertility decision, recognizing the potential endogeneity of marriage, divorce and female labor force participation rates and focusing on the under-thirty age group
– the formation of material aspirations: is parental income a good proxy, and if so what are the formative years – and what is the relationship, if any, between the intergenerational transmission of material aspirations and the intergenerational transmission of family size preferences and desired child quality?
– the theoretical potential for self-generating population cycles

Each of these is a valid area of research in its own right, and failure to establish a relationship in any one area does not provide grounds for rejection of the hypothesis in another. In addition, in keeping with the spirit of the hypothesis it is essential to differentiate carefully between desired and actual fertility in future analyses: Easterlin, following Blake (1967), maintained that underlying desires regarding family size have been fairly constant over time, with achieved fertility fluctuating as couples adjust their ideals to economic realities. Thus relative cohort size/income analyses should focus on *actual* fertility behavior, not ideal or desired. And any such analysis should make allowance for Bean and Swicegood's (1979) finding that relative income effects might be associated most strongly with a "relaxation of the need for fertility regulation", rather than with the demand for births *per se* (see Sect. 6.2.2). And finally, there is a need for further research into the macroeconomic effects of the age structure of the population: to what extent have observed 'period' effects on fertility really been attributable to changing cohort sizes?

Appendix

Subjective Measures of Relative Income

The following is a list of measures used to assess individuals' relative income in seven microlevel analyses with findings unsupportive of the Easterlin hypothesis: Crimmins-Gardner and Ewer (1978: 'CGE'); MacDonald and Rindfuss (1978: 'MR'); Olneck and Wolfe (1978: 'OW'); Thornton (1978: 'T78', 1979: 'T79'; and 1980: 'T80'); Westoff and Ryder (1977: 'WR').

"objective"
– comparisons between
  a) husband's current income and the mean income in his age-education-occupation group (T78)
  b) husband's current income and the mean income in his education group (WR, who assumed that age was also held constant, "roughly", since the respondents were all women aged 35–44)
  c) husband's occupation at marriage, currently and projected in ten years, and that of the husband's and wife's fathers (CGE)
  d) husband's income and parental head's income and welfare status – but no attempt was made to identify or distinguish between cases in which the parent was the husband's and cases in which the parent was the wife's (T80)
  e) brothers' 1973 monthly earnings difference and difference in their numbers of children (OW)

Both (a) and (b) were categorized on a five-point scale (e.g., "low, below average, average, above average, high"). No information was provided regarding the exact formulation of (c). (d) and (e) were continuous variables.

"subjective":

1) wife's response to the question "When you first got married, did you feel that you were better off or worse off financially than your parents were when you were an adolescent/growing up/16?"
   Categorical responses: "better off", "same" or "worse off".
   (used by CGE, MR, T78, and T79. CGE also asked this question with regard to current status and status in ten years. T78 also asked for a comparison between own income and that of "friends and relatives" and "past circumstances and expectations".)
2) "How well-off do you feel?" & "How do you think things will turn out in the future? Categorical responses with three to five categories, e.g., "very good", "good", "not very good". (used by T78 and T79)

T79 only:

3) Seven-point scale of responses from "terrible" to "delighted" to
   a) "How do you feel about your life as a whole?"
   b) "How do you feel about the income you and your family have?"
   c) "How do you feel about your standard of living – the things you have like housing, car, furniture, recreation and the like?"
   d) "How do you feel about the extent to which you are achieving success and getting ahead?"
4) "What about your current total family income – is that enough for you and your family to live as comfortably as you would like at this time?" Four point scale from "very comfortably" to "not at all comfortably".
5) "During the first two or three (years/months) of your (first) marriage, was your total family income enough for you and your family to live as comfortably as you wanted? at that time?" if yes – did you have most, some or few? if no: did you lack many, some or just a few?
6) Five point scale of responses ranging from "It would mean a great deal" to "I would not want it" to questions about how they would feel if they were able to... (e.g., take a vacation, buy a new refrigerator)?"

References

Abeysinghe T (1991) On testing Easterlin's hypothesis using relative cohort size as a proxy for relative income. *Journal of Population Economics* 4:53–69

Abeysinghe T (1993) Time cost, relative income and fertility in Canada. *Journal of Population Economics* 6:189–198

Ahlburg, DA (1982) The New Kuznets Cycle: A Test of the Easterlin-Wachter-Wachter Hypothesis. In: J Simon, P Lindert (eds) *Research in Population Economics, Vol. 4.* JAI Press: Greenwich, CT

Ahlburg DA (1983) Good Times, Bad Times: A Study of the Future Path of U.S. Fertility. *Social Biology* 30:17–23

Ahlburg DA (1984) Commodity Aspirations in Easterlin's Relative Income Theory of Fertility. *Social Biology* 31(3/4):201–207

Ahlburg DA (1986) A Relative Cohort Size Forecasting Model of Canadian Total Live Births. *Social Biology* 33:51–56

Ahlburg DA (1987) Investigating Fertility Fluctuations. Unpublished paper, presented at the European Population Conference, Jyväskylä, Finland

Anderson JM (1982) An Economic-Demographic Model of the U.S. Labor Market. In: J Simon, P Lindert (eds) *Research in Population Economics, Vol. 4.* JAI Press, Greenwich, CT

Artzrouni MA, Easterlin RA (1982) Birth History, Age Structure, and Post World War II Fertility in Ten Developed Countries: An Exploratory Empirical Analysis. *Genus* 38 (3–4): 81–99

Bahr SJ, Chadwick BA, Stauss JH (1975) The Effect of Relative Economic Status on Fertility. *The Journal of Marriage and the Family* 37:335–343

Baird, AJ (1987) A Note on the Easterlin Model of Fertility in Northwestern Europe and the United States: 1950–1981. *International Journal of Comparative Sociology* 23 (1–2) 57–68

Bean FD (1983) The Baby Boom and Its Explanations. *The Sociological Quarterly* 24 (Summer):353–365

Bean FD, Swicegood (1979) Intergenerational Occupational Mobility and Fertility: A Reassessment. *American Sociological Review* 44:608–618

Bean FD, Wood CH (1974) Ethnic Variations in the Relationship Between Income and Fertility. *Demography* 11 (4):629–640

Becker GS (1960) An Economic Analysis of Fertility. In: *Demographic and Economic Change in Developed Countries.* Universities-National Bureau Conference Series No. 11. Princeton University Press, Princeton

Becker GS (1976) Altruism, Egoism and Genetic Fitness: Economics and Sociology. *Journal of Economic Literature* 14:817–826

Becker GS, Lewis G (1973) On the Interaction Between the Quantity and Quality of Children. *Journal of Political Economy 81*(2): S279–S288

Behrman JR, Taubman P (1989) A Test of the Easterlin Fertility Model Using Income for Two Generations and a Comparison with the Becker Model. *Demography* 26 (1):117–123

Behrman JR, Taubman P (1990) A comparison and latent variable test of two fertile ideas. *Journal of Population Economics* 3:19–30

Ben-Porath Y (1975) First-Generation Effects on Second-Generation Fertility. *Demography* 12 (3):397–405

Berger MC (1984) Cohort Size and the Earnings Growth of Young Workers. *Industrial and Labor Relations Review* 37 (4):582–591

Berger MC (1985) The Effect of Cohort Size on Earnings Growth: A Re-examination of the Evidence. *Journal of Political Economy* 93:561–573

Bernhardt EM (1972) Fertility and Economic Status – Some Recent Findings on Differentials in Sweden. *Population Studies* 26 (2):175–184

Blake J (1967) Family Size in the 1960s – a baffling fad? *Eugenics Quarterly* 14 (1):60–74

Bliss C (1993) Life-Style and the Standard of Living, pp. 415–436. In: M. Nüssbaum, A. Sen (eds.) *The Quality of Life*. Clarendon Press, Oxford

Brady DS, Friedman RD (1947) Savings and the Income Distribution, pp. 247–265. In: *Studies in Income and Wealth, X. NBER,* New York

Brass W (1974) Perspectives in Population Prediction: illustrated by the statistics of England and Wales. *Journal of the Royal Statistical Society, Series A* 137 (4):532–583

Brown TM (1952) Habit Persistence and Lags in Consumer Behavior. *Econometrica* 20 (3):355–371

Butz WP, Ward MP (1977) The Emergence of Countercyclical U.S. Fertility. *RAND report R-1605-NIH,* June 1977, Santa Monica, CA

Butz WP, Ward MP (1979) The Emergence of Countercyclical U.S. Fertility. *American Economic Review 69* (3):318–328

Carlson E (1992) Inverted Easterlin Fertility Cycles and Kornai's 'Soft' Budget Constraint. *Population and Development Review* 18 (4):669–688

Chaudhury RH (1977) Relative Income and Fertility. *Demography* 14 (2):179–195

Chesnais J-C (1983) La notion de cycle en démographie: la fécondité post-transitionelle, est-elle cyclique? *Population* 38:361–390

Chu CTC, Lu HC (1995) Toward a General Analysis of Endogenous Easterlin Cycles. *Journal of Population Economics* 8:35–57

Congdon P (1980) Forecasting Births in Greater London: An application of the Easterlin hypothesis. *Population Studies* 34 (July):267–278

Crimmins EM, Easterlin RA, Saito Y(1991) Preference Changes Among American Youth: Family, Work and Goods Aspirations, 1976–1986. *Population and Development Review* 17 (1):115–132

Crimmins-Gardner E, Ewer PA (1978) Relative Status and Fertility, pp. 247–259. In: J. Simon (ed.) *Research in Population Economics,* Vol.1. JAI Press, Greenwich, CT

Danziger L, Neuman S (1989) Intergenerational effects on fertility: Theory and evidence from Israel. *Journal of Population Economics* 2:25–37
Day RH, Kim KH, Macunovich DJ (1989) Complex Demoeconomic Dynamics. *Journal of Population Economics* 2:139–159
de Beer J (1991) Births and Cohort Size. *Social Biology* 38 (1–2):146–153
de Cooman E, Ermisch J, Joshi H (1988) Econometric Modelling of the Birth Rate. Centre for Economic Policy Research, Discussion Paper No. 213, January 1988, London
DeFronzo J (1976) Cross-Sectional Areal Analyses of Factors Affecting Marital Fertility; Actual Versus Relative Income. *Journal of Marriage and the Family* 38:669–676
Denton FT, Spencer BG (1975) An Exploration of the Consequences of Interaction between Income Levels and Fertility Rates, Chapter 3. pp. 39–50. In: FT Denton, BG Spencer (eds) *Population and the Economy.* Lexington Books, Lexington, MA
Devaney BL (1983) An Analysis of Variations in U.S. Fertility and Female Labor Force Participation Trends. *Demography 20* (2):147–161
Devaney BL (1986) An Analysis of Variations in U.S. Fertility and Female Labor Force Participation Trends: A Reply. *Demography 23* (1):141–142
Duesenberry JS (1949) *Income, Saving and a Theory of Consumer Behavior.* Harvard University Press, Cambridge
Easterlin RA (1961) The American Baby Boom in Historical Perspective. *American Economic Review* 51 (5):869–911, and reprinted (with an additional appendix) as Occasional Paper 79 (1962), NBER
Easterlin RA (1966a) Economic-Demographic Interactions and Long Swings in Economic Growth. *American Economic Review* 56 (5):1063–1104
Easterlin RA (1966b) On the Relation of Economic Factors to Recent and Projected Fertility Changes. *Demography 3* (1):131–151
Easterlin RA (1968) *Population, Labor Force and Long Swings in Economic Growth.* Columbia University Press for NBER, New York
Easterlin RA (1969) Towards a Socioeconomic Theory of Fertility: A Survey of Recent Research on Economic Factors in American Fertility, pp. 127–156. In: SJ Behrman, L Corsa, R Freedman (eds) *Fertility and Family Planning: A World View.* University of Michigan Press, Ann Arbor
Easterlin RA (1973a) Relative Economic Status and the American Fertility Swing, pp 170–223. In: EB Sheldon (ed) *Family Economic Behavior: Problems and Prospects.* Lippincott, Philadelphia
Easterlin RA (1973b) Does Money Buy Happiness? *The Public Interest* 1973 (Winter):3–10
Easterlin RA (1976) The Conflict Between Aspirations and Resources. *Population and Development Review* 2:417– 425
Easterlin RA (1978) What Will 1984 be Like? Socioeconomic Implications of Recent Twists in Age Structure. *Demography* 15 (4):397–432
Easterlin RA (1987a) *Birth and Fortune*, 2nd edn. Basic Books New York (first published without Chapter 10 in 1980)
Easterlin RA (1987b) Easterlin Hypothesis. In: J Eatwell, M Milgate, P Newman (eds) *The New Palgrave: A Dictionary of Economics 2.* The Stockton Press, New York, pp 1–4
Easterlin RA (1987b) Fertility. In: J Eatwell, M Milgate, P Newman (eds) *The New Palgrave: A Dictionary of Economics 2*. The Stockton Press New York, pp 302–308
Easterlin RA (1995a) Preferences and Prices in Choice of Career: The Switch to Business, 1972–1987. *Journal of Economic Behavior and Organization* 27(1, June):1–34
Easterlin Richard A. (1995b) Will Raising the Income of All Increase the Happiness of All? *Journal of Economic Behavior and Organization* 27 (1, June):35–47
Easterlin RA (1996) *Growth Triumphant: The Twenty-First Century in Historical Perspective.* The University of Michigan Press, Ann Arbor
Easterlin RA, Condran GA (1976) A Note on the Recent Fertility Swing in Australia, Canada, England and Wales, and the United States, pp. 140–151. In: H Richards (ed) *Population, Factor Movements and Economic Development: Studies Presented to Brinley Thomas.* University of Wales Press, Cardiff
Easterlin RA, MacDonald C, Macunovich DJ (1990a) Retirement Prospects of the Baby Boom Generation: A Different Perspective. *The Gerontologists* 30 (6):776–783

Easterlin RA, MacDonald C, Macunovich DJ (1990b) How Well Have the American Baby Boomers Fared? Earnings and Economic Well-Being of Young Adults 1964–1987. *Journal of Population Economics* 3:277–290

Easterlin RA, Pollak RA, Wachter ML (1980) Toward a More General Economic Model of Fertility Determination: Endogenous Preferences and Natural Fertility, pp. 81–149. In: RA Easterlin (ed) *Population and Economic Change in Developing Countries.* University of Chicago Press, Chicago

Easterlin RA, Schaeffer C, Macunovich DJ (1993) Will American Baby Boomers be Less Well Off Than Their Parents? Income, Wealth and Family Circumstances over the Life Cycle. *Population and Development Review* 19 (3):497–522

Easterlin RA, Wachter ML, Wachter SM (1978a) The Changing Impact of Population Swings on the American Economy, *Proceedings of the Royal Philosophical Society* 122 (3):119–130

Easterlin RA, Wachter ML, Wachter SM (1978b) Demographic Influences on Economic Stability: The United States Experience. *Population and Development Review* 4 (1):1–21

Ermisch JF (1979) The Relevance of the 'Easterlin Hypothesis' and the 'New Home Economics' to Fertility Movements in Great Britain. *Population Studies* 33:39–58

Ermisch JF (1980) Time Costs, Aspirations and the Effect of Economic Growth on German Fertility. *Oxford Bulletin of Economics and Statistics* 43 (1980):125–143

Ermisch JF (1982) An Investigation into the Causes of Postwar Fertility Swings. In: D Eversley, W Kollmann (eds) *Population Change and Social Planning.* Edward Arnold, London

Ermisch JF (1983) *The Political Economy of Demographic Change: Causes and implications of population trends in Great Britain,* (esp. Chapter 2: Explaining and Forecasting Fertility Developments). Heinemann, London

Ermisch JF (1988) Econometric Analysis of Birth Rate Dynamics in Britain. *Journal of Human Resources* 23 (4):563–576

Ermisch JF (1990) European Women's Employment and Fertility Again. *Journal of Population Economics* 3:3–18

Evans MDR (1990) Cohort Size and Fertility Among Blacks and Whites: U.S. Cohort Born Between 1905 and 1954. *International Review of Modern Sociology* 20 (Spring):89–104

Fair RC, Dominguez K (1991) Effects of the Changing U.S. Age Distribution on Macroeconomic Equations. *American Economic Review* 81 (5):1276–1294

Fair RC, Macunovich DJ (1996) Estimated Relative Income Effects on the Labor Force Participation of Women and Men 20–24. Cowles Foundation Discussion Paper No.1116. Yale University, March

Feichtinger G, Sorger G (1989) Self-generated fertility waves in a non-linear continuous overlapping generations model. *Journal of Population Economics* 2:267–280

Feichtinger G, Dockner EJ (1990) Capital accumulation, endogenous population growth and Easterlin cycles. *Journal of Population Economics* 3:73–87

Foster A (1990) Cohort Analysis and Demographic Translation. *Population Studies* 44:287–315

Frank RH (1985) *Choosing the Right Pond.* Oxford University Press,New York

Frauenthal J, Swick K (1983) Limit Cycle Oscillations of the Human Population. *Demography* 20:285–298

Freedman DS (1963) The Relation of Economic Status to Fertility. *American Economic Review* 53:414–426

Freedman DS (1976) Fertility, Aspirations, and Resources: A Symposium on the Easterlin Hypothesis. *Population and Development Review* 2:411–415

Freedman R, Coombs L (1966) Economic Considerations in Family Growth Decisions. *Population Studies* 21:197–222

Gilks WR (1979) *The Examination of a Model of Marital Fertility Depending on Both Age and Duration of Marriage.* Unpublished M.Sc. dissertation, University of Southampton

Goodwin DD (1990) Surveying Williams College for Easterlin's Claims of Traditional Ideals. Unpubl. paper, Williams College Department of Economics, Williamstown, MA (December 9)

Hajnal J (1947) The analysis of birth statistics in the light of the recent international recovery of the birth-rate. *Population Studies* 1:137–164

Hajnal J (1950a) Births, Marriages and Reproductivity, England and Wales, 1938–1947, pp. 303–422. In: *Papers of the Royal Commission on Population, Volume II: Reports and Selected Papers of the Statistical Committee.* HMSO,London

Hajnal J (1950b) The Analysis of Birth Statistics, 1939–1943, pp. 134–177. In: *Papers of the Royal Commission on Population, Volume II: Reports and Selected Papers of the Statistical Committee*. HMSO, London

Hajnal J (1959) The Study of Fertility and Reproduction: a survey of thirty years, pp. 11–37. In: *Thirty Years of Research in Human Fertility.* Milbank Memorial Fund, New York

Hayakawa H, Venieris Y (1977) Consumer Interdependence via Reference Groups. *Journal of Political Economy* 85:599–615

Heuser R (1976) *Fertility Tables for Birth Cohorts by Color: United States, 1917–1973.* DHEW Publication No. (HRA) 76– 1152, USDHEW, NCHS, Rockville, MD, April 1976

Hill R, Klein DM (1973) Relative Economic Status and the American Fertility Swing: Introduction, pp. 167–169. In: EB Sheldon (ed) *Family Economic Behavior.* JB Lippincott, Philadelphia

Hobcraft J, Menken J, Preston S (1982) Age, Period and Cohort Effects in Demography: A Review. *Population Index* 48 (1):4–43

Howe H, Pollak RA, Wales TJ (1979) Theory and Time Series Estimation of the Quadratic Expenditure System. *Econometrica* 47 (5):1231–1247

Johnson NE, Lean S (1985) Relative Income, Race and Fertility. *Population Studies* 39:99–112

Kapteyn A (1994) The Measurement of Household Cost Functions. *Journal of Population Economics* 7:333–350

Kapteyn A, Wansbeek T, Buyze J (1980) The Dynamics of Preference Formation. *Journal of Economic Behavior and Organization* 1:123–157

Kelly WR, Cutright P (1984) Economic and Other Determinants of Annual Change in U.S. Fertility: 1917– 1976. *Social Science Research* 13:250–267

Klein LR, Rubin H (1947–1948) A Constant-Utility Index of the Cost of Living. *Review of Economic Studies* 15 (2, No.38):84–87

Kornai J (1982) *Growth, Shortage and Efficiency.* University of California Press, Berkeley

Kunz PR (1965) The Relation of Income and Fertility. *Journal of Marriage and the Family* 27:509–513

Kyrk H (1953) *The Family in the American Economy.* The University of Chicago Press, Chicago

Lee RD (1974) The Formal Dynamics of Controlled Populations and the Echo, the Boom and the Bust. *Demography* 11:563–585

Lee RD (1976) Demographic Forecasting and the Easterlin Hypothesis. *Population and Development Review* 2:459–472

Leibenstein H (1950) Bandwagon, Snob and Veblen Effects in the Theory of Consumers' Demand. *Quarterly Journal of Economics* 64:183–207

Leibenstein H (1975) The Economic Theory of Fertility Decline. *Quarterly Journal of Economics* 89 (1):1–31

Leibenstein H (1976) The Problem of Characterizing Aspirations. *Population and Development Review* 2:427–431

Leridon H (1978) Fécondité et structure démographiques: une hypothése sur l'évolution de la fécondité depuis 1940. *Population* (Mars-Avril):441–446

Lesthaeghe R, Surkyn J (1988) Cultural Dynamics and Economic Theories of Fertility Change. *Population and Development Review* 14 (1):1–45

Lindert PH (1978) American Fertility Patterns Since the Civil War, Chapter 5, pp. 137–177. In: PH Lindert (ed) *Fertility and Scarcity in America.* Princeton University Press, Princeton

Losch A (1937) Population Cycles as a Cause of Business Cycles. *Quarterly Journal of Economics* 51:649–662

MacDonald M, RA Douthitt (1992) Consumption Theories and Consumers' Assessments of Subjective Well-Being. *The Journal of Consumer Affairs* 26 (2):243–261

MacDonald MM, Rindfuss RR (1978) Relative Economic Status and Fertility: Evidence from a Cross-section. In: J Simon (ed) *Research in Population Economics, Vol. 1*. JAI Press, Greenwich, CT

Macunovich DJ (1996a) Relative Income and Price of Time: Exploring Their Effects on U.S. Fertility and Female Labor Force Participation, pp. 223–257. In: Casterline, Lee, Foote (eds) *Fertility in the United States: New patterns, new theories.* Supplement *Population and Development Review,* vol. 22

Macunovich DJ (1996b) A Review of Recent Developments in the Economics of Fertility, pp. 91–150. In: P Menchik (ed) *Household and Family Economics.* Kluwer Academic Publishers, Boston

Macunovich DJ (1997a) Why the Baby Bust Cohorts Haven't Boomed Yet: A re-examination of cohort effects on wage inequality in the U.S. Forthcoming in Ray Marshall (ed) *Can the U.S. Restore Broadly Shared Prosperity?* Proceedings of a forum on "Social and Political Implications of Major Economic Trends," sponsored by the Carnegie Corporation of New York, held in Washington, DC, May, 1997

Macunovich DJ (1997b) Discussion Comments Related to 'Social Security: How social and how secure should it be?' by Steven Sass and Robert Triest. *Social Security Reform: Links to Saving, Investment and Growth.* Proceedings of the 41st Economic Conference of the Federal Reserve Bank of Boston, June 1997

Macunovich DJ (1998) Race and Relative Income/Price of Time Effects on U.S. Fertility. *Journal of Socio-Economics* (forthcoming)

Macunovich DJ, Easterlin RA (1988) Application of Granger-Sims causality tests to monthly fertility data, 1958–1984. *Journal of Population Economics* 1:71–88

Macunovich DJ, Easterlin RA (1990) How Parents Have Coped: The Effect of Life Cycle Demographic Decisions on the Economic Status of Pre-School Aged Children, 1964–1987. *Population and Development Review* 16 (2):299– 323

Mankiw NG, Weil DN (1989) The Baby Boom, the Baby Bust, and the Housing Market. *Regional Science and Urban Economics* 19:235–258

Maxwell NL (1991) Individual and Aggregate Influences on the Age at First Birth. *Population Research and Policy Review* 10:27–46

Modigliani F (1949) Fluctuations in the Saving-Income Ratio: a problem in economic forecasting. pp. 371–443. In: *Studies in Income and Wealth, XI.* NBER, New York

Moffitt RA (1982) Postwar Fertility Cycles and the Easterlin Hypothesis: A Life-Cycle Approach. In: *Research in Population Economics, Vol 4.* JAI Press, Greenwich, CT

Murphy M (1992) The Progression to the Third Birth in Sweden, pp. 141–156. In: J Trussell, R Hankinson, J Tilton (eds) *Demographic Applications of Event History Analysis.* Clarendon Press, Oxford

Namboodiri NK (1972). Some Observations on the Economic Framework for Fertility Analysis. *Population Studies* 26:185–206

Namboodiri NK (1981) On Factors Affecting Fertility at Different Stages in the Reproduction History: An Exercise in Cohort Analysis. *Social Forces* 59 (4):1114–1129

Ní Bhrolcháin M (1992) Period Paramount? A Critique of the Cohort Approach to Fertility. *Population and Development Review* 18 (4):599–629

O'Connell M (1978) The Effect of Changing Age Distributions on Fertility: An International Comparison, pp. 233–245. In: J Simon (ed) *Research in Population Economics,* Vol.1. JAI Press, Greenwich, CT

Ohbuchi H (1982) Empirical Tests of the Chicago Model and the Easterlin Hypothesis: a case study of Japan. *Journal of Population Studies (Jinkogaku Kenkyu)* 5:8–16

Olneck MR, Wolfe BL (1978) A Note on Some Evidence on the Easterlin Hypothesis. *Journal of Political Economy* 86 (5): 953–958

Olsen RJ (1994) Fertility and the Size of the U.S. Labor Force. *The Journal of Economic Literature* 32 (1):60–100

Oppenheimer VK (1974) Life Cycle Squeeze: The Interaction of Man's Occupation and Family Life Cycles. *Demography* 11 (2):227–245

Oppenheimer VK (1976) The Easterlin Hypothesis: Another aspect of the echo to consider. *Population and Development Review* 2:433–457

Oropesa RS (1985) Subjective Relative Affluence and Expected Family Size. *Sociology and Social Research* 69:501–515

Page HJ (1977) Patterns Underlying Fertility Schedules: a decomposition by both age and marriage duration. *Population Studies* 30:85–106

Pampel FC (1993) Relative Cohort Size and Fertility: The Socio-Political Context of the Easterlin Effect. *American Sociological Review* 58 (August):496–514
Pampel FC, Peters HE (1995) The Easterlin Effect. *Annual Review of Sociology* 21:163
Pollak RA (1970) Habit Formation and Dynamic Demand Functions. *Journal of Political Economy* 78:745–763
Pollak RA (1976a) Interdependent Preferences. *American Economic Review* 66 (3):309–320
Pollak RA (1976b) Habit Formation and Long-Run Utility Functions. *Journal of Economic Theory* 13:271–297
Pollak RA, TJ Wales (1969) Estimation of the Linear Expenditure System. *Econometrica* 37 (4):611–628
Pullum TW (1980) Separating Age, Period, and Cohort Effects in White U.S. Fertility, 1920–1970. *Social Science Research* 9:225–244
Reed FW, Udry JR, Ruppert M (1975) Relative Income and Fertility: The Analysis of Individuals' Fertility in a Biracial Sample. *Journal of Marriage and the Family* (November):799–805
Rindfuss RR, Morgan SP, Swicegood G (1988) *First Births in America: Changes in the timing of parenthood.* University of California Press, Berkeley
Rutten A, Higgs R (1984) Graphic Tests of Easterlin's Hypothesis: Science or Art? In: *Research in Population Economics*, Vol.5, pp 201–212. JAI Press, Greenwich, CT
Ryder NB (1965) The Cohort as a Concept in the Study of Social Change. *American Sociological Review* 30:843– 861
Ryder NB (1980) *The Cohort Approach.* Arno Press, New York. (Originally appeared as a 1951 PhD dissertation)
Ryder NB (1983) Cohort and Period Measures of Changing Fertility. In: RA Bulatao, RD Lee (eds) *Determinants of Fertility in Developing Countries, Vol. 2*, pp 737–756. Academic Press, London
Samuelson PA (1976) An Economist's Non-linear Model of Self-generated Fertility Waves. *Population Studies* 30:243–248
Sanderson WC (1976) On Two Schools of the Economics of Fertility. *Population and Development Review 2*:469–478
Schapiro MO (1988) Socio-Economic Effects of Relative Income and Relative Cohort Size. *Social Science Research* 17:362–383
Schultz TP (1981) *The Economics of Population.* Addison-Wesley, Reading, MA, pp 150–159
Sen A (1993) Christopher Bliss: Lifestyle and the Standard of Living In: M Nussbaum, A Sen (eds) *The Quality of Life.* Clarendon Press, Oxford, pp 437–444
Serow WJ (1980) Economic Aspects of Recent Changes in Dutch Marital Fertility. *Genus* 36:189–202.
Shields MP, Tracy RL (1986) Four Themes in Fertility Research. *Southern Economic Journal* 53:201–216.
Smith DP (1981) A Reconsideration of Easterlin Cycles. *Population Studie* 35 (2):247–264
Smith DP (1986) Comment on Barbara Devaney's. An Analysis of Variations in U.S.Fertility and Female Labor Force Participation Trends. *Demography* 23 (1):137–142
Snyder D, Canarella G (1979) Births, Marriages, and Economic Cycles. *American Statistical Association, Proceedings of the Business and Economics Section, 1978*
Stevens G (1981) Social Mobility and Fertility: two effects in one. *American Sociological Review* 46:573–585
Stone R (1954) Linear Expenditure Systems and Demand Analysis: An application to the pattern of British demand. *Economic Journal* 64 (255):511–527
Sweet JA, Rindfuss RR (1983) Those Ubiquitous Fertility Trends: United States, 1945–1979. *Social Biology* 30:127–139
Sweezy A (1971) The Economic Explanation of Fertility Changes in the United States. *Population Studies* 25 (2): 255–267
Thornton A (1978) The Relationship Between Fertility and Income, Relative Income and Subjective Well-Being, pp. 261–290. In: JL Simon (ed) *Research in Population Economics, Vol.1.* JAI. Press, Greenwhich, CT
Thornton A (1979) Fertility and Income, Consumption, Aspirations and Child Quality Standards. *Demography* 16 (2):157–175

Thornton A (1980) The Influence of First Generation Fertility and Economic Status on Second Generation Fertility. *Population and Environment* 3 (1):51–72
van der Sar NL, van Praag BMS (1993) The Evaluation Question Approach: A method of measuring attitudes. *Journal of Economic Psychology* 14:183–201
van de Stadt H, Kapteyn A, van de Geer S (1985) The Relativity of Utility: Evidence from Panel Data. *The Review of Economics and Statistics* 67 (2):179–187
van Praag BMS (1968) *Individual Welfare Functions and Consumer Behavior.* North-Holland, Amsterdam
van Praag BMS (1991) Ordinal and Cardinal Utility. *Journal of Econometrics* 50:69–89
van Praag BMS (1993) The Relativity of the Welfare Concept. pp. 362–385. In: M Nussbaum, Sen A (eds) *The Quality of Life.* Clarendon Press, Oxford
van Praag BMS, Dubnoff S, van der Sar NL (1988) On the Measurement and Explanation of Standards with Respect to Income, Age and Education. *Journal of Economic Psychology* 9 (4):481–498
van Praag BMS, Kapteyn A (1973) Further Evidence on the Individual Welfare Function of Income: An Empirical Investigation in the Netherlands. *European Economic Review* 4:33–62
van Praag BMS, Kapteyn A (1994) How Sensible is the Leyden Individual Welfare Function of Income? A Reply. *European Economic Review* 38:1817–1825
Wachter ML (1975) A Time Series Fertility Equation: The Potential for a Baby Boom in the 1980s. *International Economic Review* 16:609–624
Wachter KW (1991) Elusive Cycles: Are there dynamically possible Lee-Easterlin models for U.S. births? *Population Studies* 45:109–135
Wachter KW, Lee RD (1989) U.S. Births and Limit Cycle Models. *Demography* 26 (1):99–115
Wales TJ (1971) A Generalized Linear Expenditure Model of the Demand for Non-Durable Goods in Canada. *Canadian Journal of Economics* 4 (4):471–484
Welch F (1979) Effect of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust. *Journal of Political Economy* 87 (5, pt. 2):S65–S96
Westoff CF (1981) Review of Birth and Fortune and Great Expectations. *Population and Development Review* 7 (1):131–134
Westoff CF, Ryder N (1977) *The Contraceptive Revolution.* Princeton University Press, Princeton
Wright RE. (1989 The Easterlin Hypothesis and European Fertility Rates. *Population and Development Review* 15 (1):107–122
Wright RE, Maxim PS (1987) Canadian Fertility Trends: a Further Test of the Easterlin Hypothesis. *Review of Canadian Sociology and Anthropology* 24 (3):339–357