LUND UNIVERSITY
School of Economics and Management

Master Programme in Economic Demography

The Effect of Relative Cohort Size on Fertility

Assessing the Importance of the Easterlin Hypothesis Today

Milja Norberg
milja.norberg@hotmail.com

Abstract: Although Easterlin's hypothesis of relative income has been widely supported by the majority of researchers that have tested it, these researchers have aimed to find a relationship between relative cohort size and fertility in the same time period that Easterlin examined. In this paper I argue that for the Easterlin hypothesis to be supported, it has to hold across nations as well as across time. The effect of relative cohort size on fertility is thus tested during a later time period in this paper, with the starting point derived from where most researchers on the topic have ended their scrutiny. Panel data for five industrialised nations between 1988 and 2008 have been analysed using pooled and fixed effects models. The pooled regressions show a clear negative effect of relative cohort size on fertility levels, and the mixed results from the fixed effects regressions including interactions between country and relative cohort size show both statistically and scientifically insignificant coefficients. The lack of support for the Easterlin hypothesis during this later time period indicates that the relationship between relative cohort size and fertility only was unique for the post-war period when the baby boom generation increasingly entered the labour market and fertility decreased substantially. This study discusses the possibility that a causal relationship has dissolved alternatively that the earlier relationship observed in fact was spurious due to different simultaneous processes.

Key words: Easterlin hypothesis, relative income, fertility

EKHM51
Master Thesis (15 credits ECTS)
June 2015
Supervisor: Kirk Scott
Examiner: Jonas Helgertz
Word count: 14 051

Acknowledgements

I would like to thank Kirk Scott at the Centre for Economic Demography for the discussions of the Easterlin hypothesis and feedback of my study. I am also very grateful to Frankie Videtta, David Morrow, and Nasos Alexis for their corrections of my English writing. The paper is dedicated to two people: to Frankie Videtta, who has lifted me up during this year of intense studying and always been there, and to Diego Cattolica, for his interest in others' work, major encouragement, and overall selflessness.

Table of Contents

1. Introduction

Easterlin's hypothesis of relative income has had a major impact on the literature of fertility behaviour. Not only did Easterlin challenge Becker's neoclassical economic fertility model by incorporating sociology in the economic reasoning for fertility behaviour, but he also opened up for a model that had a possibility to predict labour market outcomes and future fertility. Two broad assertions come from Easterlin's claims. The first concerns the economic disadvantage faced by the baby boom generation of the 1950s relative to their predecessors when they, as young adults, entered the labour market. Their relatively large cohort size doomed them to a life of harsher competition, and because they were economically disadvantaged in comparison to their parents' generation, fell short of their material aspirations. The second assertion follows the consequences of the relative income. The unattained aspirations will put pressure on young adults born into large cohorts to postpone family formation, and thus end up with a lower fertility rate in comparison to their parents' generation. Therefore, a self-generating mechanism of approximately 40 years is taking place where small cohorts give birth to large cohorts and vice versa (Easterlin 1980).

The contribution to the fertility literature did, since the hypothesis' first forerunner in 1961, spur a large debate with numerous scholars aiming to test the validity of Easterlin's arguments. Research spread from the initial focus on the United States to include other industrial nations in order to test whether the hypothesis would hold across nations. Although less supportive outside of Europe, the majority of the research has supported Easterlin's hypothesis of relative income (Macunovich 1998). However, the boom of research that arose with its introduction has been highly concentrated on scrutinising the same post-war period as Easterlin did, and alarmingly little research have been conducted on more recent time periods. In this paper I argue that the Easterlin hypothesis has to hold not only across nations, but also across time, if it is to be supported as a theory of fertility behaviour.

With the objective to fill the apparent time gap in the literature on Easterlin's hypothesis, I have conducted an exploratory comparative study of the United States, Finland, the Netherlands, Norway and Sweden between the years of 1988 and 2008. The research has been built upon the hypothesis as specified by Easterlin but has, because of its comparison between countries, included renowned control variables derived from the Easterlin literature. The study is breaking ground within the literature revolving the Easterlin hypothesis by using tempo-adjusted total fertility rate as the dependent variable, thus adjusting for tempo distortions that are present in the more commonly used period total fertility rate measure.

The main research question is a direct reflection of Easterlin's first assertion;

What effect does relative cohort size have on fertility?

There are two contradicting hypotheses that follow in the footstep of the main research question; either the Easterlin hypothesis is supported and we observe a positive effect of relative cohort size on fertility levels, or else it is not. The main research question is followed by two sub-questions for deeper analysis;

How do the results differ from the earlier time period studied by Easterlin? and

Why, if that is the case, do we observe a difference?

In order to answer these questions, this paper has been divided into three main parts with sub-sections. The theoretical framework presents the Easterlin hypothesis, discusses previous literature, and concludes with a combination of the two that put focus on the relevant variables. The following part revolves around the methodology of the study. It justifies the choice of research design, discusses the data and present descriptive statistics to the reader before presenting the statistical approach of the study. The later part presents and interprets the results of the study, and ends with an analytical discussion of the findings.

2. Theoretical Framework

2.1. The Easterlin Hypothesis

The Easterlin hypothesis of relative income started to take shape already in the late 1950s, when Easterlin was studying the dramatic baby boom experience in the post-war United States. His research combined with unavoidable comparisons of the opportunities of his own generation to that of his children led him to suspect that "United States might be involved in a self-generating mechanism, by which low fertility in one twenty-year period led to high fertility in the next, and vice versa" (Easterlin 1980:ix). Based upon the above-mentioned suspicion, Easterlin came to argue that generation or cohort size, i.e. the number of persons born in a particular period, has a stronger influence of shaping lives than previously believed.

Expressed simply; young adults born into small cohort sizes will experience relatively good economic conditions, which will result in higher fertility levels in comparison to young adults born into large cohorts. At the core of his claim that relative cohort size affects fertility levels lays the change in material aspirations. A deeper explanation of the concept is thus needed before we return to a demonstration of Easterlin's idea of how young couples' fertility decisions are shaped during their lifetimes.

Easterlin argued that life-style, i.e. the expectations of material standards among young couples, are unconsciously formed during childhood years. The material aspiration of a young adult is a product of the environment in which he or she grew up, which in turn was highly shaped by the income of the parents at the time. While other factors, such as peers, religion, or neighbourhood, also affect aspirations, Easterlin claimed the majority were also a product of the parents' income. Material aspirations tend to rise with each generation in industrial countries where economic development is increasing the living standards, and luxuries for a parents' generation may be perceived as necessities for their children. The minimum level of living standards that young couples perceive necessary for childbearing thus tend to increase with time, and children will be viewed as relatively less attractive than goods until the sufficient level of material aspirations is met. Although economic development also raises real income and thus makes it possible with both more children and more goods if tastes had remained the same, the raise in parents' perception of necessity levels for childbearing makes fertility behaviour subject to relative income and the impact it has on lifestyle. The relation between income and aspirations for young couples make up the relative income, which Easterlin (1980:42) defined as;

$$Relative\ income = \frac{Earnings\ potential\ of\ couple}{Material\ aspirations\ of\ couple}$$

As young couples often marry within the same social class, Easterlin argued that the parents' income of either the husband or the wife could approximate the average material aspirations, and chose to use the parents of the husband in his simplification of the definition;

$$Relative\ income = \frac{\text{Recent income experience of young man}}{\text{Past income of young man's parents}}$$

A young man's relative income to his parents' past income thus works as a link between Easterlin's two assertions when he argued that large cohorts, in comparison to small cohorts, struggle to meet their material aspirations, which leads to a lower fertility rate. In order to demonstrate the first assertion – how relative cohort size affects relative income, we will look at Easterlin's example of a young man born into a large cohort.

According to Easterlin, small cohorts will give birth to large cohorts and vice versa in a self-generating cycle. A child born into a large cohort would hence grow up with parents from a small cohort. His parents have relatively high income in comparison to their own parents, and the child grows up in an environment where high material aspirations are created. Because of the large size of his cohort, however, he is during his entire life constantly facing competition, not only when it comes to attention from parents and in school when he is a child, but especially when he as a young adult enters the labour market. At this time of life, the effect of generation size are strongest felt and shape decisions of couples that are likely to have consequences for the rest of their life-span. Large cohorts translate into a relative surplus in the labour market due to the effect of birth rate on the relative number of young people reaching working age. Correspondingly, the effect of the labour market surplus on earnings and employment results in unfavourable economic life chances faced by large cohorts relative to small. The young man born into a large cohort thus experiences relatively high unemployment, has a hard time to find a job with a good wage, and struggles to advance on the career ladder. Although real income in relative terms has increased substantially over time, his income relative to his material aspirations will be low, which will affect his wellbeing and behaviour.

The above demonstration of Easterlin's first assertion opens up for possibilities to examine the second assertion; that relative income affects fertility levels. Easterlin empirically displays a general higher degree of mental stress due to failure to meet expectations. Doubt of own ability spurs resentment towards both others and self, and crime, suicide and political alienation due to psychological stress are shown to be more prevalent among young adults born into large cohorts than small. Failure to meet expectations tends to put strains on relationships as well, increasing the divorce rate and relative number of illegitimate births. A low relative income pressures young adults to sacrifice family formation while seeking opportunities to increase their economic earnings. According to Easterlin, husbands' insufficient incomes of large cohorts are a strong reason for increased female labour force participation. This reasoning is based upon the assumption that socially constructed gender roles in the family remain, and that while men are faced with pressure to be good providers, women are firstly judged on their accomplishments as mothers. Although the increase over time has been striking, Easterlin claimed that the unusual rise in young women's labour force participation since the 60s is a sign of declining relative incomes as large cohorts reach young adulthood. Easterlin's data shows that while the increase in female labour force participation is seen as a continuous phenomenon, the rapid increase in the 60s and 70s is a strong contrast to the almost zero growth rate two decades before. Easterlin's explanation is that for young couples who are not able to live the life they want to due to low relative income levels of the men, one strong possibility to improve the couples' economic situations is by increasing the women's time at work away from home. Putting off childbearing or to have fewer children in order for the wives to return to work sooner, are consequences to these decisions. When Easterlin presented his hypothesis, he recognised that much had been done to change traditional sex-role images as well as that real signs of change were present, but used empirical evidence in order to show that no fundamental shifts had taken place among the population when it came to the roles within the family. The women remained expected to drop out of the labour force to care for the children to a much larger degree than men, and the expected jobs of men and women were still the traditional occupations and also the most sought after. Hence, Easterlin in his attempt to explain human behaviour argued that traditional family roles remained, and primarily focused on male relative income as the driving force of fertility.

With young adults born into large cohorts feeling pressure to sacrifice family formation, a lower fertility level will be present. According to Easterlin, these factors have created self-generating cycles of approximately 40 years. Small cohorts, due to their relatively high incomes, are more likely to have children at a young adult age, thus giving birth to relatively large cohorts. For large cohorts, the opposite is true. Although generation size is likely to always have had a limited impact on life outcomes, Easterlin emphasised that the importance of generation size has increased significantly with new conditions that arose in American Society since the Second World War. The federal government implemented policies that severely restricted immigration and maintained a growing employment level, thus altering the American labour supply and demand conditions which, as Easterlin argued, "resulted in a new relation between population and the economy" (Easterlin 1980:5).

2.2. Previous Literature

The Easterlin hypothesis has been subject to heavy scrutiny since its introduction to the fertility discussion. The dominant figure in reviews of the large literature on the Easterlin hypothesis is Macunovich, who in her impressive review from 1998 critically assesses seventy-six published analyses of the hypothesis. As a former student of Easterlin, Macunovich's words are interesting from the point of view that she, possibly more than other scholars, are more informed of what Easterlin was testing in his hypothesis. Her perhaps biased critique towards research that have failed to support the hypothesis is in fact that many scholars seem to have misinterpreted the main points with the hypothesis, and hence conducted studies that do not resemble Easterlin's ideas. Although the support of the Easterlin hypothesis varies in degrees across countries, Macunovich finds an unambiguous support for the impact of relative income on fertility.

Especially studies of the Easterlin hypothesis within North America have tended to be supportive. Ahlburg (1982) used data from the United States to find significant coefficients in sub-periods already since the 1920s, but attributed the failure of finding significant coefficients during the entire period since the beginning of 1900s to the change in labour supply due to immigration. Apart from Ahlburg, there have been several other scholars that have supported the Easterlin hypothesis on a macro level. Although there are studies focused on North American data that find unsupportive results, e.g. Rutten and Higgs' (1984) visual analysis of the data used by Easterlin himself but at different time periods and scales, authors like Macunovich (1998) argue that they tend to draw conclusions not in favour of the Easterlin hypothesis by falsely representing Easterlin's arguments.

The most critical of the Easterlin hypothesis has been Ermisch, who has tested the hypothesis on countries outside North America five times (1979) (1980) (1982) (1983) (1988). Ermisch started with visual analyses the way Easterlin had, and continued with OLS regressions. He initially found that relative economic status could be one of many factors influencing fertility but rejected the hypothesis by and large both then and in later studies due to the only weak support by the evidence he found. Major setbacks, as Macunovich highlighted, were his failure to construct the relative income variable based on age-specific rates to represent the relationship between old and young workers' incomes. In one study where he followed Butz and Ward's (1979) use of age-specific relative income, he did in fact provide evidence that relative cohort size had a significant effect on fertility of young women, but continued his further studies using a faulty relative income measure and unambiguously argued against the Easterlin hypothesis. He was also criticised for excluding important external factors, and when he did by including real house prices and women's relative wages, received critique for their endogeneity. Other authors testing the Easterlin hypothesis outside of North America have been more positive, but largely varied in their choice of methods and countries studied. These studies have ranged from macro-studies of socialist Eastern European countries, visual inspection of total fertility rates in Japan and its correlation to both relative cohort size and relative income, as well as an analysis based upon Israeli micro data, showing supportive results that nevertheless were difficult to interpret with unstable coefficients and multicollinearity problems (Carlson 1992)(Ohbuchi 1982)(Danziger and Neuman 1989) (Macunovich 1998).

However, if the Easterlin hypothesis is to be revolutionary in explaining fertility behaviour, the hypothesis should be significant on a cross-country level. Only a relatively limited amount of cross-country studies have been made in comparison to single-country analyses, also here varying between visual inspection, OLS, and Granger Causality analytical techniques. What they share in common is that they all test the relationship between relative cohort size and fertility. When comparing their results country by country, Macunovich found similarities across the studies, which mostly tended to show strong supporting results for e.g. the United States, Australia, New Zealand and Wales, an inverse relationship for Germany, and an absence of a relationship for Portugal and Spain. Although comparing the studies may be hazardous due to the different ways their results are presented, Macunovich argue that countries such as Finland, Norway and the Netherlands, which in the cross-country comparisons have shown a possibility of a relationship or non-significant ones, can perhaps be placed among the countries with supporting results. Differences among the studies include for example Baird's (1987) double use of relative measures by including both the age ratio and relative income, and O'Connell's (1978) different indicator of cohort size picking up the leading and lagging differences between cohorts rather than between large and small cohorts.

Pampel also stands out from the crowd by including measures of collectivism in his article from 1993. I especially want to highlight his study as Macunovich clearly argues that one major setback of many cross-country studies is their exclusion of significant factors that can explain differences between countries. Pampel incorporated a number of variables in order to control for institutional differences in a comparison of 18 developed countries. The rationale behind his study was, just like Macunovich noted, that while support for the Easterlin hypothesis had been strong in the United States, support in European countries had proven weak. According to Pampel, differences in institutional structures of social protection play a large role in influencing relative economic status, and he argued that "societal institutions promoting collective responsibility for living standards and solidaristic policies of social protection cushion the harmful impact of large cohort size on economic well-being" (Pampel 1993:499). In other words, he showed that countries with strong collectivist support systems tend to have a more limited effect of relative cohort size on fertility, as the negative economic situations associated with large cohort sizes and their oversupply of workers at young adult age may be limited due to policies that keep unemployment to a minimum and guarantee jobs. The knowledge of the existence of these policies may also impose a sense of security for what risks the future might bring, and reduce the postponement of marriage and childbearing for these groups. Social benefits can also contribute in providing financial security for the unemployed, as well as governmental subsidies aimed at certain vulnerable groups in society can do. Pampel argued that just as institutions of social protection can shape young adults' economic status in a direct way, the institutional environment of a society could indirectly shape the interpretations of relative economic status. As nations vary in their degree of social protection, differences across countries are likely to show varying support for the Easterlin hypothesis as childbearing decisions among young couples in countries that are highly committed to social protection may be less conditioned on the relative economic status. Indeed, Pampel finds that countries with low levels of collectivism, such as the United States, have a much stronger effect of relative cohort size on fertility than countries such as Finland and Sweden with high collectivism. Although Pampel's study has been highlighted as revolutionary in many ways, with his OLS-model results confirmed by granger causality analysis, his inclusion of institutional settings-variables stand out as a single rarity among studies of the Easterlin hypothesis.

A perhaps more debated variable among the scholars of the Easterlin hypothesis is the controversy of including female labour force participation as a variable. As explained in the previous section, Easterlin's predictions were based upon the effect of birth cohort size on male relative income, excluding young women's economic contribution due to the prevailing family roles in society. Already in 1976, Oppenheimer was among the first scholars to criticise the sole use of male relative income, and a long debate has followed since. Firstly, Oppenheimer criticised Easterlin for comparing relative economic status between father and son, while estimating the changes in their market positions over time though a general – and not male-specific - unemployment rate. She also noted that when Easterlin talked about material aspirations as derived from a comparison of own income relative to parents' income at the time of their own childhood years, it is family income rather than only father's income that is considered. Oppenheimer argued that there has been a change in the relative economic status of women, which have had consequences on wives' economic contributions to their families, and that their incomes also must be counted. Oppenheimer suggests that the decline of male economic position relative to parents initiated a response of increasing female labour force participation as a compensation to increase income in the family which in turn led to young couples more easily reaching their preferred material life style. While this goes hand in hand with the Easterlin hypothesis, Oppenheimer argues that the following generation's young adults had become more economically disadvantaged in comparison to their parents, and sacrifices in childbearing and marriage and an increase in wives' labour force participation, became necessary to reach an affluent life-style. This pattern has continued, and each new cohort has had to increase wives' labour force participation and reduce early fertility, which Oppenheimer argues is what has happened. Wives' economic contribution in the family has thus increased by generation and become an important factor in the formation of a family.

Oppenheimer's contribution in the Easterlin hypothesis discussion has been large, and also made other scholars propose that a more accurate determinant of fertility behaviour would be to include females in the relative cohort size variable as to highlight how females also compete for jobs at young adult ages (Baird 1987) (Pampel 1993). Her contribution also spurred an inclusion of a female labour force participation variable among scholars, arguing that as wives' participation in the labour force can compensate for low male incomes among large cohorts, the measure indicates women's position in society and may have an impact on relative cohorts' effect on fertility. As women's position in society varies between countries, it is of extra importance in cross-country analyses (Pampel 1993).

Inclusion of the variable may however cause endogeneity bias due to the relationship between female labour force participation and fertility. Fertility decisions are highly possible to be determined in conjunction with the factors we argue to be the main determinants of fertility, such as female labour force participation. In other words, one factor may not cause the other, but rather be jointly determined. The problem with endogeneity bias in regressions is that the relations are hard to identify which makes them difficult to estimate, and inconsistent OLS-estimators will remain a problem even if we can identify the relations (McNown 2000). Many researchers looking at fertility have nevertheless included the variable, and while some have been criticised for using it without providing any alternative regressions where the variable is excluded, others have aimed to lag the participation rate or used instrumental variable procedures (Pampel 1993) (Macunovich 1998). According to McNown (2000), these attempts to reduce the endogeneity bias do not go far enough, as he argues that "[t]he entire system of variables involved in aggregate fertility models is subject to rampant endogeneity" due to them all being outcomes of interdependent decisions made by young couples (McNown 2000:7). Hence, the complexity of endogeneity calls for caution when variables such as female labour force participation are included, but they may nevertheless prove important due to the structural changes of the roles of women in society.

2.3. The Focus

The two previous sections have discussed Easterlin's hypothesis from his own perspective and presented previous literature that have aimed to validate the hypothesis. The large scope of the Easterlin hypothesis opens up for numerous different research possibilities. This calls for a clarification and closer look at the gaps in the already existing literature, as well as the major variables necessary.

Easterlin clearly emphasised that immigration and business cycles affect the supply and demand for labour at entry levels, which is why he argued that relative cohort size increased in significance when the United States implemented policies that restricted immigration and maintained a growing level of employment. Many scholars have included the unemployment rate as a control variable to capture the impact of economic cycles, which highlights the variable's importance. Migration, on the other hand, is indirectly controlled for in the most common way of measuring relative cohort size, and thus tends to have been excluded as a separate variable in previous research. The relative cohort size measure work as a suitable proxy to relative income, and has been the key independent variable used by all cross-country studies as well as many single-country studies. In the majority of studies, the measure has followed Easterlin’s example, with a ratio of the old male adults aged 30- 64 to young male adults aged 15-29. The large debate of changing female roles in society has however spurred many authors to include females in the ratio, as their income play a larger role in total family income than previously acknowledged, and thus cannot be neglected to have an impact on relative income and fertility behaviour. Women’s changing roles have also spurred the debate of whether female labour force participation rate should be included as a variable or not. While Easterlin in the 1980s was eager to explain that the improvement of gender roles had not structurally changed the traditional family roles, it is today hard to neglect that female labour force participation is not solely an outcome of a deteriorating male relative income. That the trend has continued until today is a fact, but not much has been written or empirically tested of the Easterlin hypothesis during the last decade. A new look at the validity of the Easterlin hypothesis is in order not only to scrutinise the changing gender roles, but also to highlight years that have not yet been looked at. Including females in the relative cohort size measure would emphasise wives’ economic contribution in the family, while female labour force participation would control for differences in women’s contribution across countries. Because of the potential endogeneity bias, however, analysis of the Easterlin hypothesis ought to provide alternatives where female labour force participation is not included and possibly causing biased outcomes. Some of the more important variables highlighted in Macunovich’s review, but largely neglected by other scholars, are Pampel’s inclusion of variables that control for institutional differences. The varying degrees of social protection across countries, and the strong effect social benefits and protection policies may have on relative income levels and fertility behaviour make such control variables crucial in a cross-country analysis.

An immediate distinction between my research and previous research on the Easterlin hypothesis is the measure of childbearing. Easterlin himself argued that in order to test the relative income hypothesis, the total fertility rate must be used as a dependent variable as it reflects changes in marital fertility as well as marriage behaviour. These changes, according to Easterlin, are dominated by the behaviour of women under thirty years of age, whose children born account for three quarters of all children. The total fertility rate in a year is in his book defined as “the total number of children that a hypothetical woman would have borne if she had gone through her reproductive life having children at the average rate of childbearing actually prevailing in that year at each age from fifteen to forty-four" (Easterlin 1980:48). This definition has been imitated in most other studies of the Easterlin hypothesis. In fact, the conventional indicator of fertility in a year, used by the majority of scholars focusing on demographic issues, is the period total fertility rate (Potancoková, Sobotka, and Philipov 2008). There are, however, some strong tempo-distortions that arise with the use of this measure, which have to be explained to justify the use of the new fertility measure.

The period total fertility rate is affected by two components; the quantum component, which is the level of fertility, and the tempo component, which is the timing of childbirth. It is the tempo component that lately has become of interest due to the structural postponement of births in most European countries since the early 1970s, which has shifted the mean age of childbearing. The tempo effect distortion that arises with this postponement transition is that although the number of births over the life course remains constant, there will be a decline in the number of children born during the early womanhood years, hence depressing the period total fertility rate (Potancoková, Sobotka, and Philipov 2008). The risk for tempo-distortions were introduced by Ryder already in the late 1950s when he demonstrated that the discrepancy between the period total fertility rate and the cohort completed fertility rate depended in size upon how fast the mean age of childbearing was changing (Bongaarts and Sobotka 2012). First in proposing a method to measure the quantum of fertility in a year without including the tempo-effect discussed above was Bongaarts and Feeney in 1998, when they used the age of mother and birth order of child to better indicate the average number of births per woman, defining tempo distortions as "an inflation or deflation of the period TFR when the period (instead of the cohort) mean age at childbearing changes" (Bongaarts and Sobotka 2012:91). In 2012, Bongaarts and Sobotka highlighted the usefulness of a new variant of the measure introduced by Bongaarts and Feeney; tempo- and parity-adjusted total fertility rate, which provided more stable values by also controlling for the parity composition of the female population. They used it in a comparison to other total fertility measures to estimate the "role of declines in tempo and parity composition distortions in the recent rise in the conventional total fertility rate in Europe", arguing that the period total fertility rate creates misinterpretations of the levels of fertility while the old tempo-adjusted total fertility rate is neither controlling for the parity distribution nor the instability from year to year

(Bongaarts and Sobotka 2012:84)[1]. The relatively new measure of total fertility rate has been included in the Human Fertility Database, making the measure easily accessible without complex calculations required by the researcher, thus allowing for more detailed analyses of fertility trends than before. That no other studies of the Easterlin hypothesis have yet included the measure of adjusted total fertility rate is hence not surprising, but the major setbacks with period total fertility nevertheless makes a focus on the adjusted total fertility rate necessary for more accuracy.

3. Methodology

3.1. Research Strategy and Design

The previous section indicates that there is a strong need to narrow the scope of a study based upon the Easterlin hypothesis. The hypothesis ranges across decades and includes numerous interesting aspects, hence there is no surprise that both micro and macro studies have been carried out, as mentioned in the previous literature. While the main theoretical steps would be to examine either what impact birth cohort sizes have on relative income or what impact relative income has on fertility behaviour, the Easterlin hypothesis also touches on questions of how material aspirations are formed and how factors such as divorce, suicide and crime are affected by relative income and affect fertility behaviour. Qualitative research strategies could help us to achieve an in-depth understanding of some of these factors. However, a quantitative approach is more suitable in this case for a number of reasons. As an alternative to Becker's neoclassical economic fertility model, the important role of the Easterlin hypothesis lies in whether it can be generalised to a large extent. With the new time period aimed to fill a gap in the literature on the hypothesis, this study has been limited to test the main point across time and nations, and a cross-country comparison of fertility behaviour over time requires the use of aggregated data. In line with other quantitative cross-country comparisons of the hypothesis, the study has used relative cohort size as a proxy for birth cohort as well as relative income in order to test the impact on fertility.

[1] See Bongaarts and Sobotka's article "A Demographic Explanation for the Recent Rise in European Fertility" for a full explanation of how the tempo- and parity-adjusted total fertility rate is calculated.

The secondary data used are collected on a regular basis on samples of the population in a repeated cross sectional design. It is highly unlikely that the sample remains the same, as it would in a longitudinal study, and we therefore cannot fully address the direction of cause and effect. Aggregated data collected over a long period of time, however, makes it possible to chart change that can indicate causal relationships on country levels. Ideal in establishing a causal relationship would be to do an experimental design, but because of rare opportunities to do so, control variables are required in cross-country comparisons to account for differences across the countries studied that may affect the results (Bryman 2012). Before turning to the statistical approach of the study, the two following sections will present the variables of interest and the descriptive statistics of the data.

3.2. Data

The analysis covers five democratic welfare states between 1988 and 2008. The countries chosen are based upon Pampel's study of the Easterlin hypothesis from 1993, where he chose 18 industrial countries to compare the effect of relative cohort size on fertility. Limited data availability of the dependent fertility variable has been the main determinant of limiting the number of countries studied, and also affected the years of study. However, the main restriction to the years included has been the idea that the hypothesis should hold also when tested at time periods outside Easterlin's scope. Country specific data of Finland, the Netherlands, Norway, Sweden and the United States have been collected from the United Nations World Population Prospects 2012, the Organisation for Economic Co-operation and Development, and the Human Fertility Database.

While Easterlin hypothesised that relative cohort size could explain the levels of fertility within a country, there is a strong possibility that the relationship also holds across countries. Major differences between countries must therefore be controlled for, and allowing certain control variables to determine fertility by interacting countries with relative cohort sizes can help to identify the effect of the Easterlin hypothesis across countries. Internal validity bias is relatively limited but it should be acknowledged that the countries chosen are far from a representative sample of all industrialised countries in the world, and that other variables than the ones chosen may affect the relationship between cohort size and tempo-adjusted total fertility rate. Limited time resources in combination with a willingness to simplify have determined the most crucial control variables based upon the Easterlin hypothesis and previous research on the topic. A big advantage with the official data sources used is that the data is relatively harmonised in order to ensure country-comparisons. Official sources often use country specific data that follow international guidelines, which specify how data are to be collected as well as how central concepts should be defined. The use of highly renowned databases also limits the risk that society will consider the data sources used in the research as unethical.

The dependent variable is, contrary to most studies of the Easterlin Hypothesis using period total fertility rate, the tempo-adjusted total fertility rate. The rate indicates the average number of children per woman specified by birth order of the child and age of the mother. The data, because of its relatively new way of measuring fertility with reduced tempo-distortions, is rather limited but can be accessed through the Human Fertility Database. The database encourages researchers to use other indicators of fertility than the period total fertility rate, and uses officially registered births by calendar year, the age or cohort of the mother, and biological birth order in their indicators. The same measures have been used for all countries in order to produce uniform data. The data available at the Human Fertility Database is limited to countries with reliable population estimates ranging all reproductive ages and where there is complete birth registrations made by official statistical agencies. This is the main reason for the limitation to only five countries, but it also increases the reliability of the data.

Relative cohort size has been used as a key independent variable, again following the example of Easterlin's claim that relative cohort is a suitable proxy for both relative income and birth cohort. However, while Easterlin argued that relative cohort should be measured as the ratio of male population of 30-64 to 15-29, I have chosen to follow Oppenheimer's (1976) argument of changing gender roles and wives' increasing economic contribution to the family income, and therefore included the female population in the ratio used. Age-specific data have been derived from the UN World Population Prospects 2012 and merged to the correct age groups before calculating the ratio as population aged 30-64 divided by population aged 15-29. A high ratio will display that the younger population is relatively small to the older population, thus symbolising a small cohort entering the labour market with relatively good opportunities of obtaining a high relative income. Again, the measure of population is rather accurate for the countries chosen because of their high standard of official registration.

However, for each country past trends of mortality and international migration is accounted for together with fertility, where international migration being the hardest to project. Lack of information of age-distribution of migrant flows have resulted in that the researchers at the United Nations World Population Prospect 2012 have used models based upon assumed migration flows to distribute the net number of migrants by age-group (United Nations 2014).

Unemployment rate is included as a control variable for fluctuations in the labour market due to economic cycles. The variable is important for the Easterlin hypothesis' basic idea that relative cohort size only became a significant variable in the United States after the Second World War when migration was limited and supply of jobs became relatively even, thus changing the supply and demand for young workers. While migration is accounted for in the measurement of relative cohort size, fluctuations in the labour market still need to be controlled for. Data on unemployment rates, measured as total unemployment for the working population aged 15 to 64, are derived from Organisation for Economic Co-operation and Development on labour statistics. Although argued to facilitate international comparisons, the national surveys on which the data are based may differ in sample and time aspects. Unemployment is measured slightly differently. Of the five countries, there are differences in the population included, ranging from pertaining to all registered in the country to only including private households. In the Swedish survey from 2005, also Swedish people employed abroad are included in the working population (OECD notes 2014). The values also vary from being based upon monthly, quarterly or semestrial estimates.

The variables of social security spending and family allowance spending are based upon Pampel's (1993) inclusion of numerous variables that accounted for differences in social institutions across countries. Although Pampel included both dynamic and stable variables in his regressions, and ranked the countries after the merged scores of each estimated variable in each country, I have chosen to include the two variables that are easily attainable and theoretically remain strong. Social security spending controls for country variations in public spending on benefits for social protection and reflect the commitment to redistribute resources with social purposes. Low-income households, elderly, sick, disabled, young, and unemployed citizens are usually targeted through these policies, which have the power to cushion poor conditions in the labour market for young adults. The data on public social security spending as a percentage of GDP is collected from the Organisation for Economic Cooperation and Development, as is the data on public family allowance spending, also measured as a percentage of GDP. Family allowance spending exclusively targets parents with children, thus differing from social security spending. Both variables reflect countries' commitment to social protection for economically vulnerable young adults, and facilitate family formation and childbearing though financial support. The effects of relative cohort size on fertility are likely to be weaker in countries with strong commitment to social protection, and the variables are therefore included to control for cross-country differences. The variable would ideally measure social protection spending directed at young adults, as policies directed at the young are more likely to have an impact on fertility decisions among young couples, but because of the unavailability of information on how much is spent on the young, the social security spending variable rather remains a valuable indicator of nations' commitment to social protection.

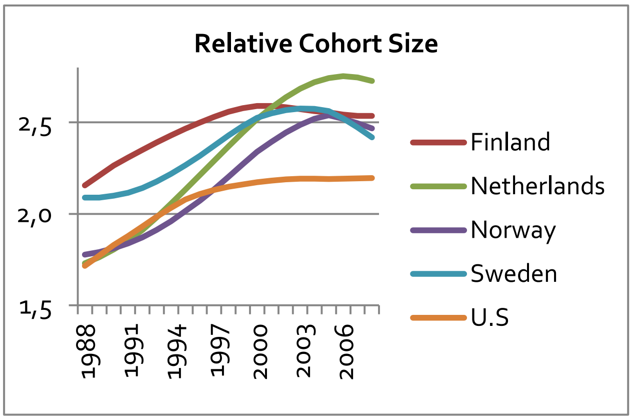
The final variable is the much-debated female labour force participation rate included to acknowledge wives' increased economic status and its likelihood to affect the relationship between relative cohort size and fertility. The data on female labour force participation is collected from the Organisation for Economic Co-operation and Development on labour force statistics. It measures the annual female labour force participation rate as a percentage of the population, and concentrates on the ages between 15 and 29 in order to limit the ages to young adulthood years. As discussed in previous sections, the variable remains subject to endogeneity bias. The variable has been included in half of the regressions in order to reflect how the economic roles of women have changed across years and nations, and excluded in half of them to provide less biased estimates.

3.3. Descriptive Statistics

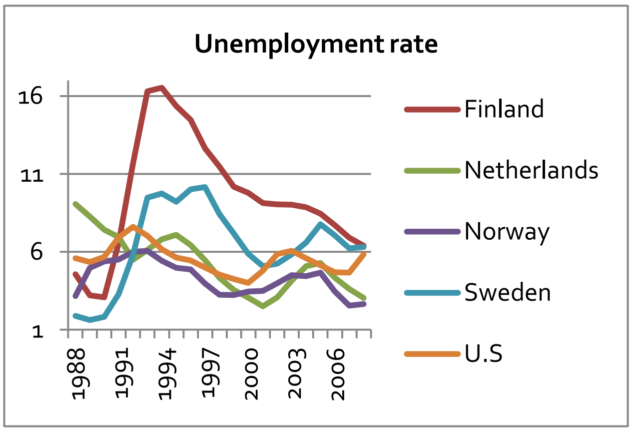
The following table presents summary statistics of the collected data:

Variable	Number of observations	Mean	Standard deviation	Min	Max
Tempo-adjusted TFR	105	2,007	0,172	1,70	2,40
Relative cohort size	105	2,279	0,276	1,72	2,75
Unemployment rate	105	6,220	2,948	1,62	16,53
Female labour force participation rate	105	70,941	6,407	50,64	82,46
Social security spending	105	23,104	5,515	12,80	35,50
Family allowance spending	105	2,456	1,225	0,41	4,80

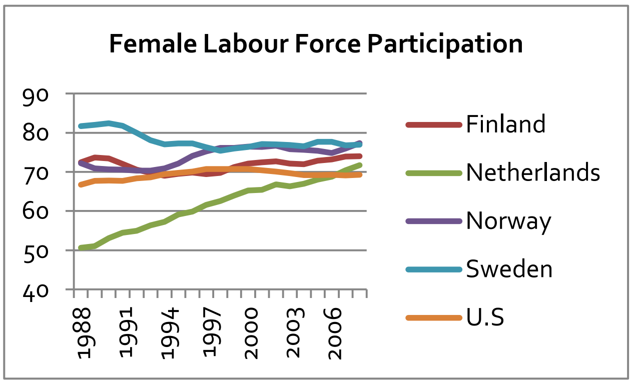
The collected data has not been altered substantially, but rather presents the values derived from each variable's data source. No outliers were detected. The relative cohort size, our key independent variable, ranges between 1.72 and 2.75. As the variable is measured as the population aged 30-64 divided by the population aged 15-29, our values indicate that at a minimum, the older population is 1.7 times more populous than the younger. The higher the ratio, the more populous is the older population to the younger, and the smaller is the relative cohort size of the young adults. In other words, a high ratio implies that the cohorts of young adult age are relatively small compared to if the ratio is low, and according to the Easterlin hypothesis we can assume that the small cohort have a high relative income. A closer look at the line diagram of the relative cohort size data below can show us further information of the change over time for the separate countries. For all countries, relative cohort size has increased between 1985 and 2008, however, it is possible to see slight fluctuations in which Sweden, Norway and the Netherlands had static or declining numbers during the first years of study and all countries tend to show declining or static numbers around the year of 2000.



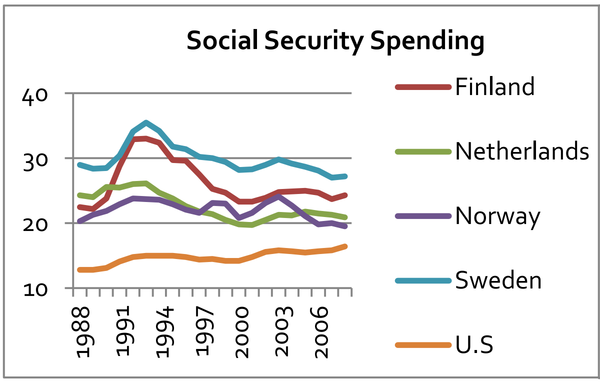
The unemployment rate has a large variation across countries and years, and varies between 1.62 and 16.53. The line diagram for unemployment rate shows a major increase during the financial crisis of the early 1990s, where Finland's unemployment rate reaches above 16% of the working population. On an average, the unemployment rate over the years has fluctuated near 6% for all counties together.

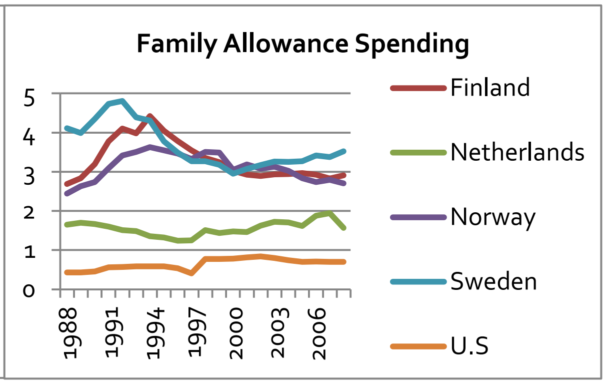


Female labour force participation has perhaps the largest variation with a standard deviation of 6.4 and ranging between 50.64 and 82.46 per cent of the female population. The diagram to the right shows us that the Netherlands on its own strongly contribute to lowering the average to just above 70.9%, but has remarkably increased its' female labour force participation between 1988 and 2008, indicating a strong change in women's economic contribution to the family. Apart from the Netherlands, the other countries display relatively stagnant rates across the years. There is no strong indicator that the participation rates for young women have fluctuated since the 1980s in accordance with Easterlin's hypothesis.

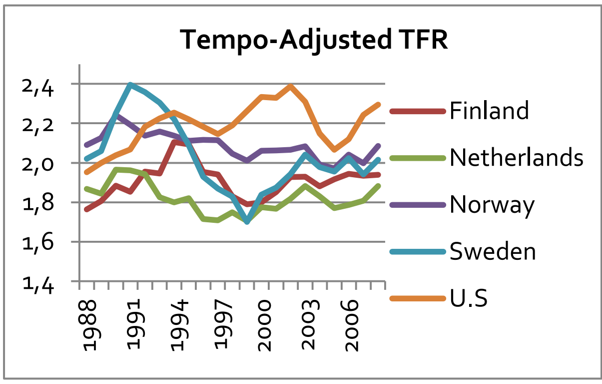


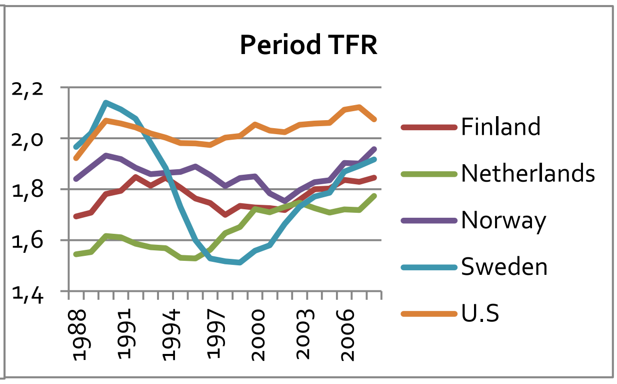
The summary description also shows us the government spending on social security and family allowance as a percentage of GDP for the different countries. Social security spending as well as family allowance spending is on average the highest for Sweden and lowest for the United States, as the diagram below displays. As a minimum, the United States spent 12.8% of GDP on social security and 0.41% on family allowance, while Sweden at maximum reached 35.5% on social security spending and 4.8% on family allowance spending. The variation between the countries indicates different governmental commitment to social protection, and emphasises the importance of controlling for them in the regressions.





Our dependent variable, the tempo-adjusted total fertility rate, is also interesting to take a closer look at. As the summary statistics shows us, it varies between 1.7 children per woman to 2.4 with relatively little variance as can be seen by the standard deviation for the variable. Although not included in the data, a line diagram of period total fertility rate has been constructed in order to display the differences between the two measures. As we can see, the period total fertility rate looks relatively similar to the tempo-adjusted total fertility rate in the overall shapes of the curves. However, the values vary and are on average higher for the tempo-adjusted total fertility rate, while it also shows fluctuations in more detail. In spite of the tempo-adjusted total fertility rate for Sweden, the rates show only small fluctuations and remain stable during the 20-year period.





Concerns for normal distribution of the residuals were tested through kernel density estimations, and the data deviate somewhat from the normal curve, especially the family allowance spending variable that reveals skewness but no heavy or light tails.

3.4. Statistical Approach

In order to analyse the secondary data, multivariate analysis has been used to include the crucial control variables as well as the changes in the relationship between adjusted total fertility rate and relative cohort size over time. The data is structured as balanced long panel data over cross sections and has been analysed using the software programme STATA12 for Windows.

Regressions on panel data can capture variations over both country and time. As each time period of data is dependent on previous time periods, the standard errors need to be adjusted. This is not the case of cross-sections, and thus panel data is more complicated and requires estimation methods and models that are richer. Nevertheless, with the use of panel data that includes the variables of unemployment rate, female labour force participation rate, social security spending and family allowance spending, it is possible to control for country-specific effects that may be related to geographical, historical or political contexts. While there are many different linear models to be used for panel data, the pooled regressions and the fixed effects models are the ones emphasised in this paper. The fundamental distinction between different panel models are between fixed effects models and random effects models, where a random effects model assumes that the individual-specific effects are completely random and thus uncorrelated with the regressors whereas a fixed effects model does not. The Hausman test has implied that the panel data of this study is more suitable for fixed effects than random effects models. Pooled regressions and fixed effects models can more accurately help to answer the main research question of this study.

A pooled regression is similar to an Ordinary least squares (OLS) model, except that we add an extra dummy for each year. The effect of relative cohort size (RCS) on tempo-adjusted total fertility rate (TFR) can be specified in an econometric model by moving from model specification 1 to model specification 2 below.

Model specification 1:

$$TFR_{it} = \beta_0 + \beta_1 RCS_{it} + \beta_2 unemployment_{it} + \beta_3 flfp_{it} + \beta_4 social_{it} + \beta_5 family_{it} + V_t + \alpha_i + u_{it}$$

In this first model, the tempo-adjusted total fertility rate in country *i* at time *t* equals the intercept plus the effects of the relative cohort size (RCS), unemployment rate, female labour force participation rate (flfp), social security spending (social), and family allowance spending (family) in country *i* at time *t*. Adding to this comes a whole range of other unobserved factors, some which are solely time dependent (V_t), some that are solely country dependent and do not vary with time (α_i), and some which are characteristic factors that vary both with time and country (u_{it}). In the second model specification, the pooled OLS model, dummy variables have been created for each year, excluding the first year to avoid the dummy variable trap. For reasons of simplification, in this basic demonstration of the second model, the middle control variables and dummy variables have been excluded.

Model specification 2:

$$TFR_{it} = \beta_0 + \beta_1 RCS_{it} + \cdots + \beta_5 family_{it} + \gamma_1 \delta_{2t} + \cdots + \gamma_{T-1} \delta_{Tt} + \alpha_i + u_{it}$$

The pooled OLS regression allows us to see how relative cohort size affects the tempo-adjusted total fertility rate over time. It is however important to remember that the coefficients yield an overall measure rather than a separate one for each year. The actual error we see in our regression is thereby now including both the country dependent factors (α_i) and characteristic factors (μ_{it}). What a pooled OLS regression does is lump all observations together and estimates them as if they had all belonged to the same country by creating a line of best fit. Thereby it also assumes that the regressors are exogenous and that the countries pooled are homogenous. The observations across different countries and across different times are treated as if they were randomly sampled observations. The problem with this model is that for OLS estimates to be consistent it requires that the covariance of the error with the independent variables is equal to zero. Even if we control for certain independent variables in our model, this is likely not to be the case as there is a strong risk that the country dependent factors are correlated to the one or more of the independent variables. The covariance of the country specific factors with the independent variables will not be equal to zero, and the pooled regression estimates will be inconsistent.

Because of this problem of country dependent factors, which we call unobserved heterogeneity, I will also use fixed effects models to estimate the relationship between relative cohort size and tempo-adjusted total fertility rate. In a fixed effects model, we calculate the time averaged equation of the pooled regression specification, and then subtract the time averaged equation from the original specification. As the country dependent variable (α_i) does not vary with time, the time averaged country dependent variable remains in its original form, and is thus removed from the equation when the time averaged equation is subtracted from the original specification. The unobserved heterogeneity will be removed and the regression will be consistent.

Time averaged equation: $\overline{TFR}_i = \beta_1 \overline{RCS}_i + \cdots + \beta_5 \overline{family}_i + \gamma_1 \delta_{2t} + \cdots + \gamma_{T-1} \delta_{Tt} + \alpha_i + \bar{u}_i$

Time averaged equation subtracted from original equation:

$$TFR_{it} - \overline{TFR}_i = \beta_0 + \beta_1 (RCS_{it} - \overline{RCS}_i) + \cdots + \beta_5 \left(family_{it} - \overline{family}_i\right) + \gamma_1 \delta_{2t} + \cdots + \gamma_{T-1} \delta_{Tt} + (\alpha_i - \alpha_i) + (u_{it} - \bar{u}_i)$$

Model specification 3:

$$\widetilde{TFR}_{it} = \beta_0 + \beta_1 \widetilde{RCS}_{it} + \cdots + \beta_5 \widetilde{family}_{it} + \gamma_1 \delta_{2t} + \cdots + \gamma_{T-1} \delta_{Tt} + \tilde{u}_{it}$$

The problem with a fixed effects model is that it removes anything that is time constant, and we thus cannot evaluate the effect of time-constant variables on the dependent variable.

In one fixed effects model, I will also include interaction effects between each country and relative cohort size. By generating dummy variables for each country except the United States, which then becomes the reference category while we avoid the dummy trap, and then interacting each variable with relative cohort size, it is possible to differentiate how the effect of relative cohort size on fertility differs between the five countries.

Model specification 4:

$$\widetilde{TFR}_{it} = \beta_0 + \beta_1\widetilde{RCS}_{it} + \cdots + \beta_5\widetilde{family}_{it} + \gamma_1\delta_{2t} + \cdots + \gamma_{T-1}\delta_{Tt} + \beta_6(RCS \times Finland) + \beta_7(RCS \times the\ Netherlands) + \beta_8(RCS \times Norway) + \beta_9(RCS \times Sweden) + \tilde{u}_{it}$$

The analysis of the effect of relative cohort size on fertility is, for each model, carried out both with and without the female labour force participation rate variable. Including the variable is in one way necessary in order to recognise the mediating effects that it may have on the relative cohort size effects, but the potential endogeneity on the other hand makes it important to test regressions where the variable is excluded.

Both pooled OLS regressions and fixed effects models will be used to analyse the relationship between relative cohort size and tempo-adjusted total fertility rate. Period effects have been controlled for through the use of year dummies. All independent variables have been lagged with one year to account for the fact that birth does not occur until approximately a year after fertility decisions, which are likely to be strongly affected by the conditions we are controlling for in the independent variables. The concern of reliability, that the results found are repeatable and the measures consistent, is tested with the use of Cronbach's alpha test. All estimates will also be considered in relation to the related significance level. The significance level, which is the probability of rejecting the null hypothesis given that it is true, must be higher than the p-value, which is the probability of random chance to explain the result, for a coefficient to be statistically significant. A p-value of 0.05 thus indicates that there are only 5% probability that the results are due to chance, and thus 95% probability that the results are not due to chance (Cameron and Trivedi 2009). Significance levels of 0.05 and 0.01 have been used in this study to indicate whether the coefficients are statistically significant or not.

4. Analysis

4.1. Results Pooled Regression

Two models, one with and one without the female labour force participation rate variable (FLFP), were tested through pooled OLS regressions of the five countries between 1988 and 2008 (see Table 1). The R^2, which measures how close the observed data are to the fitted regression line, is 74.3% in the FLFP-model and 55.9% in the model without, indicating that the FLFP-model better explains the variation of data. However, there is a risk that the R^2 can be misleading, as it cannot determine if estimates and predictions are subject to bias. A closer look at the residual plots helped to confirm that the models fit the data relatively well, with randomly dispersed residual plots around each variable's horizontal axis. Multicollinearity among the independent variables in the models was also tested for, by using the Variance Inflation Factor (vif) command. While values over 10 indicate strong multicollinearity, the highest value was 7.64 for family allowance spending in the model including female labour force participation rate as a variable. The models have been adjusted for heteroskedaticicity with the robust command, as variance that is spread unequally makes the estimators inefficient and may lead to an overestimation of t- and f-statistics.

Table 1: Pooled OLS regression

Measure	FLFP included		FLFP not included	
	coeff (β)	Standard Error	coeff (β)	Standard Error
Relative cohort size	-0.563**	0.069	-0.570**	0.074
Unemployment rate	0.011*	0.005	-0.000	0.006
Social security spending	-0.010*	0.004	-0.022**	0.005
Family allowance spending	0.002	0.021	0.115**	0.022
Female labour force participation rate	0.019**	0.003		
Constant	1.915**	0.225	3.223**	0.150
R^2	0.743		0.559	

Note: Level of significance, * = p<0.05, ** = p<0.01

Analysis of the pooled data shows that the relationship between relative cohort size and tempo-adjusted total fertility rate is significant but negative for the five countries between 1988 and 2008, both with and without controlling for female labour force participation. According to the Easterlin hypothesis, relatively smaller cohorts tend to, because of their higher relative income, have a higher fertility rate. The relative cohort size variable in the dataset is measured so that the higher the value is, the smaller the cohort is. According to the Easterlin hypothesis, we would thus expect that an increase in the relative cohort size variable would have a positive effect on tempo-adjusted total fertility rate. However, the results from the pooled regressions show that for each additional unit increase in the relative cohort size variable, the tempo-adjusted total fertility rate decreases - with 0.563 children when female labour force participation is controlled for, and 0.570 children when it is not. Both coefficients were highly significant at the 0.01 level.

Although the control variables are of little importance when it comes to interpreting the data, the big variation in family allowance spending between the two models requires a comment[2]. The model including female labour force participation rate had a coefficient of 0.002 and a non-significant p-value as high as 0.90, while the model without the variable had a coefficient of 0.118 and was significant to the 0.01 level. We do, however, have to remember that all control variables are crucial to control for differences between countries, and the effect we see lumps all countries together, and thus averages the experiences. As discussed previously, the relationship in the pooled OLS model may therefore be affected by heterogeneity between countries. We turn to the fixed effects models in the following sub-section to look at countries individually.

4.2. Results Fixed Effects

Table 2: Fixed Effects Model

Measure	**FLFP included**		**FLFP not included**	
	coeff (β)	Standard Error	coeff (β)	Standard Error
Relative cohort size	-0.460*	0.127	-0.267	0.114
Unemployment rate	0.002	0.017	-0.006	0.012
Social security spending	0.001	0.010	-0.003	0.010
Family allowance spending	0.150	0.081	0.193	0.071
Female labour force participation rate	0.013	0.009		
Constant	1.529*	0.515	2.141**	0.260
R^2	0.650		0.632	

Note: Level of significance, * = p<0.05, ** = p<0.01

Table 2 shows fixed effects models, with and without female labour force participation rate (FLFP) included as a variable. In these models, the unobserved heterogeneity has been removed in the specification, making the regressions more consistent than the pooled OLS regressions. Country dependent factors previously included in the error term have thus been removed. The models were adjusted for heteroskedaticity and autocorrelation through the vce(cluster country) command.

[2] As the main research question of this study is looking at the effect of relative cohort size on fertility, the control variables have been not interpreted in the results sections. They are included in the regressions to control for differences between countries, but their own effects on fertility in each country are not of interest to this study.

In the fixed effect model including female labour force participation rate, we can see that the relative cohort size variable remains negative with on average 0.46 less children per unit increase in the relative cohort size variable. This value is less negative than the relative cohort size value in the pooled OLS regression, but significant only to the 0.05 instead of the 0.01 level. For the model without female labour force participation included, the negative value is even smaller, but has become insignificant. While the constant for both models are significant to the 0.05 level, the all of the independent variables apart from relative cohort size are non-significant. The R^2s have converged to 0.650 and 0.632 from previous 0.743 and 0.559. For both the pooled regressions and fixed effects models, a more negative effect of relative cohort size on fertility seems to be more statistically significant.

Table 3: Fixed Effects Model with Country and Cohort Interaction

Measure	FLFP included		FLFP not included	
	coeff (β)	Standard Error	coeff (β)	Standard Error
Relative cohort size for the United States	0.523	0.297	0.676	0.384
Unemployment rate	0.012	0.008	-0.005	0.009
Social security spending	-0.009	0.006	-0.006	0.011
Family allowance spending	0.140	0.056	0.146	0.069
Female labour force participation rate	0.033	0.016		
Interaction Finland	-0.249	0.235	-0.262	0.214
Interaction Netherlands	-1.074*	0.266	-0.717	0.261
Interaction Norway	-0.826**	0.136	-0.741**	0.157
Interaction Sweden	-0.400	0.375	-0.763*	0.243
Constant	-0.485	1.323	1.461	0.580
R^2	0.803		0.766	

Note: Level of significance, * = p<0.05, ** = p<0.01

Including interactions between each dummy variable for country and the relative cohort size variable allows the effect to vary between countries, displayed by Table 3. The models make it possible to see if the relationship between relative cohort size and tempo-adjusted total fertility rate is different depending on which country we look at. Table 3 shows us that the R^2 has strongly increased for both models, while the relative cohort size effects have changed further. The interaction effects have to be interpreted in relation to a reference country, which in this study was chosen to be the United States as it was the country where Easterlin himself carried out his visual study. Table 3 shows us that if the coefficients are correctly estimated, the relative cohort size effect is positive for the United States while all country interaction effects are negative. The calculation of relative cohort size effect for the European countries will be explained below, and an analysis of the results will follow in the upcoming discussion section.

In the model where female labour force participation is included, the relative cohort size effect for the United States is 0.523. For each unit increase of the relative cohort size variable, the tempo-adjusted total fertility rate increases with 0.523 children. Table 4 helps us to structure the following values for each country. In Finland, this effect is 0.249 children less than in the United States; 0.274 children in total. The Netherlands has the largest negative value of 1.074 children less than in the United States – in total decreasing the tempo-adjusted total fertility rate with 0.551 children for each unit increase in the relative cohort size variable. Norway also has a large negative interaction effect of 0.826 children less than in the United States, resulting in a decreasing tempo-adjusted total fertility rate coefficient of 0.303. Lastly, the interaction effect for Sweden is 0.4 less than United States, and manages to keep a small but positive coefficient of 0.123 increase in the fertility rate for each additional unit of relative cohort size.

Table 4: FLFP included

Country	Interaction coeff (β)	Relative cohort size effect
United States	0 (ref.)	0.523
Finland	-0.249	0.274
Netherlands	-1.074*	-0.551
Norway	-0.826**	-0.303
Sweden	-0.400	0.123

Note: Relative cohort size effect calculated by the base effect of the reference country minus the interaction effect coefficient.

As previously explained, an increase in the relative cohort size variable symbolise a decrease in the cohort size of young workers relative to old workers. According to the Easterlin hypothesis, an increase in the relative cohort size variable is expected to result in a higher tempo-adjusted total fertility rate. A positive coefficient, which is in line with the hypothesis, is seen for the United States, Finland, and Sweden. On the other hand, the only statistically significant coefficients are for the Netherlands and Norway, who show negative relative cohort size effects.

Table 5: FLFP not included		
Country	Interaction coeff (β)	Relative cohort size effect
United States	0 (ref.)	0.676
Finland	-0.262	0.414
Netherlands	-0.717	-0.041
Norway	-0.741**	-0.065
Sweden	-0.763*	-0.087
Note: Relative cohort size effect calculated by the base effect of the reference country minus the interaction effect coefficient.		

In the model where the female labour force participation rate is not included, the effect of relative cohort size on tempo-adjusted total fertility rate is higher for the United States than in the previous model, reaching an average of 0.676 children. Once again, all European countries have negative interaction effects, but in this model only Finland, with an interaction effect of 0.262 lower than the United States, manages to keep a positive effect of relative cohort size on tempo-adjusted total fertility rate. The Netherlands, Norway and Sweden, on the other hand, all show negative relative cohort size effects. The negative effects they display, 0.041 for the Netherlands, 0.065 for Norway and 0.087 for Sweden, are less negative than in the model where the female labour force participation rate variable was included. However, Sweden has shifted from being positive and statistically insignificant, to the most negative effect and statistically significant. While the Netherlands have lost its significance, Norway keeps being negative and significant to a 0.01 level.

As with the previous models, the more negative effects tend to be more statistically significant.

4.3. Discussion

The Easterlin hypothesis argues that the size of a birth cohort has a significant impact on fertility behaviour. In comparison to small cohorts, large cohorts are faced with unfavourable economic conditions due to a harsher competition for entry-level jobs and career opportunities. They will have a harder time in achieving their material aspirations created during their childhood years, and will experience pressure to postpone childbearing until their aspirations are met. Young wives will increasingly participate in the labour force to contribute to the insufficient income of the husbands, further postponing childbearing. Large cohorts will thus give birth to small cohorts, who will not only grow up in scarcity with opposite conditions and outcomes, but who will also have lower relative material aspirations due to their parents' unfavourable economic situation. A self-generating circle of approximately 40 years, with large cohorts giving birth to small cohorts and vice versa, is what Easterlin argued to have started after the Second World War when the U.S. government policy stabilised the demand for and restricted the supply of labour.

The above-simplified repetition of the Easterlin hypothesis is crucial to remind us of how the relationship between relative cohort size and fertility according to Easterlin has created a self-generating cycle. In order to analyse whether the Easterlin hypothesis hold across countries and time, this research has aimed to be precise with Easterlin's main points, while walking in the footsteps of previous literature on the hypothesis. In order to avoid the critique pointed at unsupportive scholars such as Ermisch, this study has been careful to use the measurement of relative cohort size as specified by Easterlin. According to the hypothesis, we would therefore expect a positive effect of relative cohort size on fertility, indicating that the smaller the size of the cohort, the higher the rate of fertility.

However, if we return to the results of the pooled regressions, fixed effects models, and interactions, we do find them unsupportive of the Easterlin hypothesis. Both pooled regressions, which in an ideal world of the Easterlin hypothesis would show a positive effect of relative cohort size on fertility across industrialised nations, instead display statistically significant negative results. In the more consistent fixed effects models, the results continue to be negative, but only the model including the female labour force rate variable remains statistically significant. The fixed effects model including interactions between relative cohort size and countries makes it possible to analyse the effect within each country, and contains the results of most interest for this study. If we assume that the coefficients are correctly estimated, the results vary, both across nations, and depending on whether we include the possibly endogenous female labour force participation rate variable or not. If we do, we find a positive effect for the reference country the United States, and negative interaction effects of the European countries. When the interaction effects are deducted from the relative cohort size effect of the United States, Finland and Sweden show small positive effects of relative cohort size on fertility, while the Netherlands and Norway display negative effects. In the model where the possibly endogenous female labour force participation variable is excluded, the positive effect of the United States is larger, but only Finland remains with a positive effect among the European countries. These results are in line with previous findings by other scholars - the Unites States tend to show a positive effect, while European countries instead tend to show negative ones.

However, the results from the interaction models call for a closer look at the statistical significance of the estimates. In fact, only the interaction effects for the Netherlands and Norway are statistically significant in the model where female labour force participation is included, and in similar fashion, it is only Norway and Sweden that remain statistically significant in the model where female labour force participation is not. In both models, it is the more negative interaction effects that are statistically significant. The fact that the United States is statistically insignificant while it is the reference country makes interpretations of the interaction effects more difficult. For both models, there is a 15,3% risk that the observed relative cohort size coefficients for the United States have arisen purely by chance. Although this is a high risk that is far from accepted in econometric modelling, the risks are even greater for the European countries that, if assumed correct, are showing positive results. Finland and Sweden show risks of above 34% in the model where female labour force participation is included. In the model where the participation rate is not included, Finland alone shows a risk of above 27%. The other European countries, which show negative effects of relative cohort size on fertility, in both models, show risks of less than 6%. The statistical significance indicates that the positive effects of relative cohort size on fertility are more likely to be wrongly estimated than the negative effects. Therefore, we cannot rule out the possibility that the positive effects, to a great extent, risk being negative.

This discussion of statistical significance has to be put in relation to scientific, or economic, significance. A dominant scholar on the scientific significance literature is McClosky (2009), who strongly argues that statistical significance is far from sufficient to prove results to be relevant in practice. If we return to the results of this study, we thus have to acknowledge the change of cohort size required for a scientific effect to occur. One unit increase in the relative cohort size variable symbolises a doubling of the population aged 30-64 compared to the population aged 15-29. In other words, the whole size of the younger population would have to increase in the older population for the relative cohort size variable to increase with only one unit. The one unit change thus resembles an immense decrease of the young cohort size to the older, and it would require a very small relative cohort size in comparison to the older in order to see any real difference in fertility levels. The largest negative relative cohort size statistically significant effect we see, -0.551 for the Netherlands in the model where female labour force participation rate is included, becomes scientifically insignificant when interpreted as a decrease of the average number of children per woman by approximately half a child, for each doubling of the older population to the younger. For a real change in fertility levels to occur, a not yet observed change in the relative cohort size would be required. Hence, the results can be considered scientifically insignificant. Interestingly, when previous literature on the Easterlin hypothesis is examined from the same perspective of scientific significance, their results, are faced with the same dilemma of scientific insignificance.

The results from the pooled regressions, fixed effects models, and interaction effects indicate that there is no strong support of the Easterlin hypothesis between the time period of 1988 and 2008. The findings contradict the Easterlin hypothesis, as well as the majority of scholars who have tested it. If we assume that the unsupportive results found between 1988 and 2008 are correct, the major question that arises is why these results differ from the period studied by Easterlin. To answer this question, the lack of relationship in the later period has to be put in the context of the earlier periods following the baby boom after the Second World War.

The relationship between relative cohort size and fertility as examined by the majority of scholars unambiguously includes the post-war era, as the graph from Pampel's study in 1993 displays below.

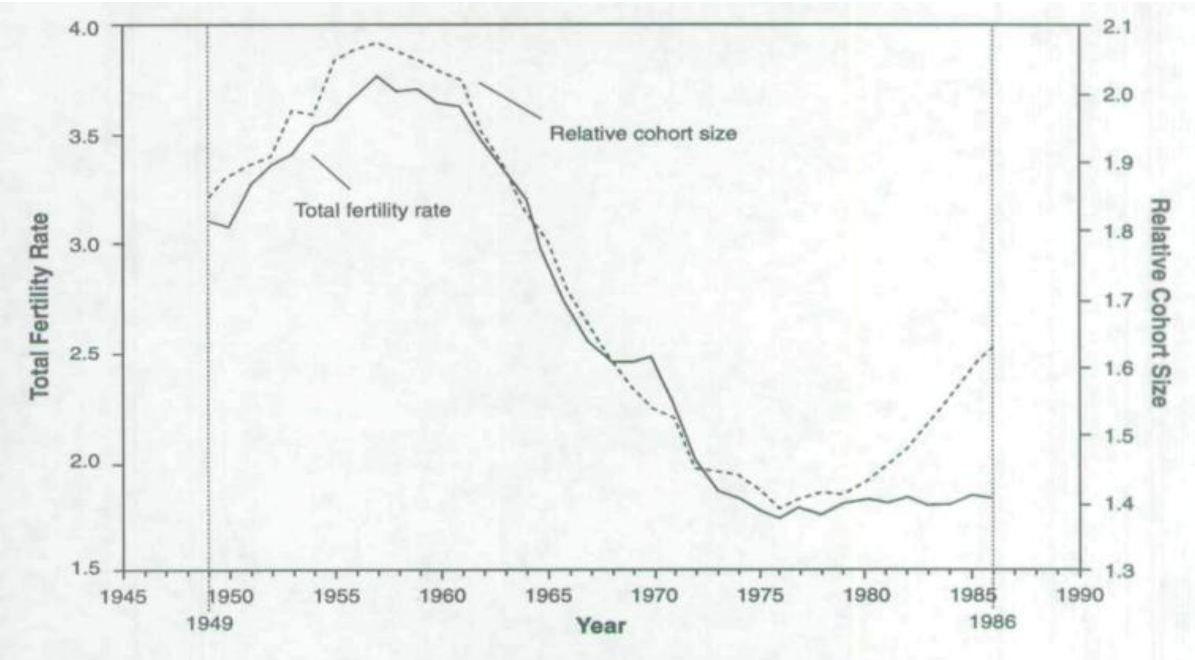


Figure 1. Total Fertility Rate and Relative Cohort Size by Year: United States, 1949 to 1986

Note: Graph taken from Fred C. Pampel's article "Relative Cohort Size and Fertility: The Socio-Political Context of the Easterlin Effect" from 1993.

The relationship is visually apparent from the 1950s. In the late 1950s, the relative cohort size variable is declining as the baby boom generation of the post-war era is entering the labour market. We also see a decline in total fertility rate. However, when the relative cohort size variable starts to increase again in the 1980s, there is no indicator that the total fertility rate does the same. Instead, the fertility rate stays relatively stagnant after the 1980s, which is confirmed by the line diagrams of both tempo-adjusted and period fertility rates in the descriptive statistics, which, apart from Sweden, show no trend in fluctuations after the 1980s. Revolutionary for this study is that it examines the hypothesis beginning from the 1980s and thus excludes the time period where the Easterlin hypothesis can be visually supported. Instead, I argue that if the Easterlin hypothesis is to be supported, it has to hold across any time period. There are a few alternatives to why the effect of relative cohort size on fertility appears positive pre-1980s and insignificant post-1980s. Firstly, the relationship during the early period may have been causal, but dissolved during the later period. Secondly, there may never have been a causal relationship, and the relationship observed by Easterlin and other scholars during the early period were rather spurious, with separate processes that caused the variables to decline simultaneously.

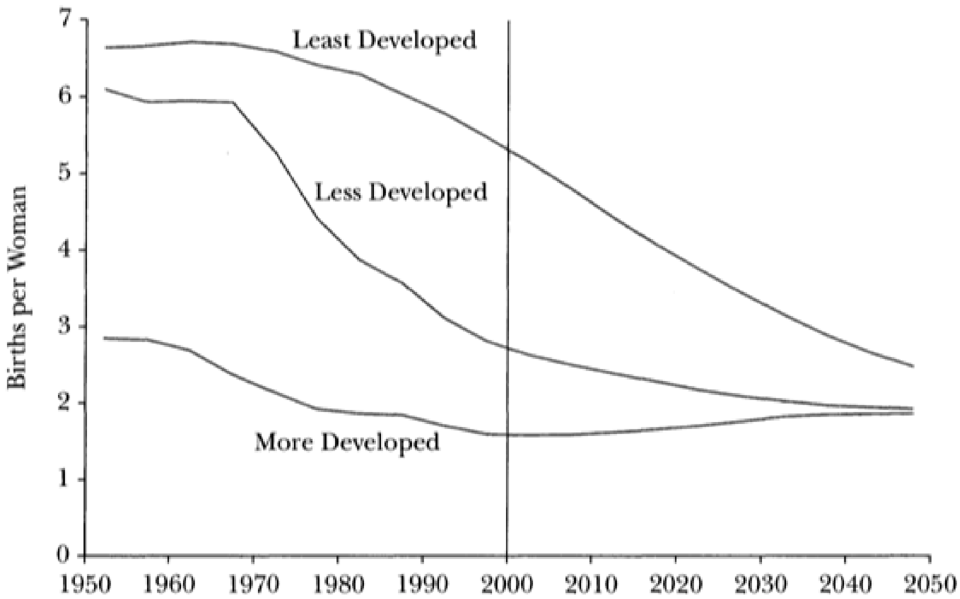
If we turn to the first alternative, it can be argued that a causal relationship existed during the time period Easterlin and other scholars observed, but has seized to exist at later time periods. In other words, if there indeed was a self-generating mechanism created by the conditions of post-world war America, it has decreased in significance on both a statistical and scientific level. Extensive literature has tried to explain the decrease in fertility and why fertility has remained low post 1980s. Rather than arguing that fertility responds in a linear way to a change in the relative cohort size variable, it can be argued that when fertility reaches a certain level or conditions change, fertility can start to behave in a non-linear way. There is a consensus in the literature that once fertility reaches low levels, a so called "trap" may occur where it will be difficult to increase fertility levels again. The change in the demographic regime from high to low fertility levels may create a self-reinforcing process that keeps fertility low, and challenges Easterlin's idea that fertility will increase when small cohort sizes enter young adult age. However, Lutz, Skirbekk and Testa (2006) include Easterlin's ideas of material aspirations as one out of three possible mechanisms that constitute this self-reinforcing low fertility trap (appendix 1).

The first mechanism is demographic and reflects how the number of births (together with mortality and migration) influences the age-structure of the population, which later affects the number of births from the period fertility. The number of births, or crude birth rate if we divide it by the population, is a function of the age-structure in combination and past period fertility, with the latter affected by cohort fertility and the timing of fertility discussed earlier as a reason for using tempo-adjusted total fertility rate, rather than period total fertility rate as a measure for fertility. Cohort size, in turn, is affected by social norms such as how many children constitute an ideal family size (Lutz, Skirbekk and Testa 2006). These social norms make up the second mechanism; the sociological one, which is assumed to affect cohort fertility. The low fertility trap hypothesis claims that declines in actual fertility have affected the ideal family size through a change in social norms. Young couples especially are influenced by the experience around them, which has seen a decline in the number of children (ibid). The low fertility trap hypothesis emphasises the family situation of couples only slightly older as having a stronger influence, as opposed to couple's parents' past economic situation, which was argued by Easterlin. The economic rationale is thus the third mechanism included in the low-fertility trap hypothesis, and is directly drawn from Easterlin's idea of couples' income relative to material aspirations formed during childhood years affecting the timing of family formation. An increasing mean age of childbearing is an effect of these social and economic factors combined, which is a reason for why tempo-adjusted total fertility rate had to be used in this study.

Lutz, Skirbekk and Testa (2006) argue that while material aspirations have continued to increase, opinion surveys have documented that young people are faced with an increasingly pessimistic economic outlook. An ageing population force changes in nations' social security systems, where young generations are more negatively affected by the necessary cuts than older generations who experience more gradual cuts. In addition, they argue that "rapid population ageing may also result in lower productivity and consequently in a globalised economy, less investment and lower economic growth in the future", which, together with deterioration of benefits for younger generations, trigger a negative perception of their economic status (Lutz et al 2006:14). This argument can also be strengthened through looking closer at female labour force participation as argued by Oppenheimer (1976), which indicated a trend in participation, which perhaps to a start followed Easterlin's theory, but in today's standards have come to symbolise a structural change. As explained in the theoretical framework of this paper, Oppenheimer argued for a change that we can refer to as a female labour force participation trap - in which the increase of participation of one generation creates a higher threshold for the following generation in order to reach their material aspirations, and further participation of wives will be necessary to reach the life style seen as a pre-requisite for family formation. With increasing participation, a change in the economic contribution of wives within the family has also improved. Such a change is likely to have helped the transformation of fertility behaviour in combination with other factors such as the low fertility trap, which together may have dissolved the assumptions of the self-generating mechanism as argued by Easterlin.

The second alternative is that the relationship observed by Easterlin and many of his followers between the 1950s and 1980s, in fact was spurious. The decline in fertility and the decline in the relative cohort size variable may have been caused by separate processes that cause the curves to follow similar patterns. As the relative cohort size variable is measured as a ratio of older to younger workers, time will cause the variable to fluctuate. The fertility variable, however, will not automatically fluctuate due to its measurement. The fertility variable thus has to be put in a larger context over a greater time period.

The graph below by Lee (2003) shows how fertility has decreased since the 1950s and is projected to decrease for different groups of countries.



Source: Historical and Middle Series forecasts are taken from United Nations (2003).

While we see a more rapid decrease among less and least developed countries, the decrease is also present for more developed countries - where the five countries of this study are included. More importantly, we see a general decrease in fertility for all countries, and not only welfare states[3]. Fluctuations in fertility levels are not visible, and the strong decline in fertility since the post-war period indicates a trend on a global level that is not dependent on cohort size fluctuations. The structural decline in fertility is defined by Lee (2003) and other scholars as the second demographic transition - where more developed countries experience beyond replacement levels of fertility, while less and least developed countries encounter substantial declines in the number of children born per woman. Other scholars argue that the second demographic transition is merely a continuation of the first one, which for developed countries already started in the beginning of the 19th century with declining mortality (Lee 2003). The drop in fertility during the period studied by Easterlin may simply have been a continuation of the fertility decline argued to have been caused by modernisation and resulting in the second demographic transition when fertility reached very low levels in many more developed countries. As Becker argued in his fertility theory, modernisation caused children to become more costly in comparison to consumption goods. Put in a historical context, technological advancement in combination with physical and human capital increased the productivity of labour, and thus the value of time (Galor and Weil 1996). Lee (2003:174) links the above argument to the development until today; "Rising incomes have shifted consumption demand toward nonagricultural goods and services, for which educated labor is a more important input. A rise in the return to education then leads to increased investments in education. Overall, these patterns have several effects: children become more expensive, their economic contributions are diminished by school time and educated parents have higher value of time, which raises the opportunity costs of childrearing".

Thus, the decline in fertility between the 1950s and 1980s may have been caused by factors related to modernisation and the demographic transition, rather than caused by an increasingly large cohort size entering the labour market in bad economic conditions. The lack of relationship between the variables post 1980 indicates that there is a likelihood that the decline in the relative cohort size variable in earlier studies just happened to occur at the same time as a structural decline in fertility, and the correlated events may have been mistaken for a causal relationship. With no other study confirming a continuation of the self-generating cycle

[3] The majority of researchers studying the Easterlin hypothesis have focused on industrial (i.e. more developed) countries.

predicted by Easterlin, where the relationship with fertility also continues when the relative cohort size variable is increasing, the results of this study from 1988 to 2008 indicates that the Easterlin hypothesis is based entirely upon a specific period in time and cannot be generalised as a continuous phenomena[4].

5. Summary conclusion

Easterlin's hypothesis of relative income has had a major impact on the literature of fertility behaviour. By suggesting that relative cohort size has an impact on relative income, which will play a strong role in fertility behaviour, Easterlin's hypothesis would be able to predict future labour market outcomes, as well as fertility outcomes. The large potential contribution spurred numerous scholars to test the Easterlin hypothesis, the majority of them confirming the relationship between relative cohort size and fertility. Nevertheless, the results from this study contradict the results from previous research.

In an exploratory country-comparative research, five industrialised countries have been examined between the years of 1988 and 2008 in order to investigate the effect of relative cohort size on fertility, in a later period than Easterlin and his followers. First, this paper has shown that the effect of relative cohort size on fertility differs between the countries studied, but due to both statistical and scientific insignificance, the effects can be disregarded. Thus, the Easterlin hypothesis cannot be supported. Second, the time period of 1988 to 2008 shows different results from supportive studies of earlier time periods. However, while contradicting results would show a clear negative instead of positive effect, the results of the latter time period indicate no significant effect at all. Third, and last, this paper has attempted to explain possible reasons to the difference in results between the two time periods. Put in relation to previous findings of the post-war period, it has discussed two major alternatives to the limited support of the more recent time period. The first one argues that even if relative cohort size strongly contributed to fertility behaviour when the baby boom generation in the United States entered the labour market in a harsh economic environment and thus experienced pressure to put off having children, the relationship has ceased to exist. While it is likely that relative

[4] The visual increase in fertility from 1949 to 1957 in Pampel's graph shows the post-war baby boom. Improved conditions after the war facilitated family formation that had been surpressed during the war.

income still matters, young couples may postpone family formation due to an increasing importance of women to participate in the labour force in order for the couples to meet their material aspirations. Also, to have fewer children later has become more socially acceptable, which is argued to have created a low-fertility trap. The second alternative instead argues that the causal relationship between relative cohort size and fertility may never have existed. The relationship observed by Easterlin and many other scholars is likely to have been supportive of the relationship only because of the unique time period they examined, but instead of the relationship being causal, it may have been spurious. While relative cohort size will fluctuate with time, fertility tends to have decreased substantially between the 1950s and 1980s, but then remained low until today. The decrease in fertility during this time period may very well be a consequence of modernisation, which happened to take place at the same time as the post-war baby boom generation started to enter the labour market. Thus, there are no significant indicators that the drop in fertility was a consequence of the relative cohort size. In either case, Easterlin's hypothesis of a self-generating mechanism has failed to occur. The fact that the Easterlin hypothesis is failing to reflect the fertility behaviour of today indicates an insignificant influence of relative cohort size to shape life outcomes. To predict labour market outcomes and fertility behaviour based upon the size of a cohort is therefore not to be recommended.

6. References

Articles and Books

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

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.

Bongaarts. J. and Feeney, G. (1998), “On the quantum and tempo of fertility”. *Population and Development Review.* 24 (2): 271-291.

Bongaarts. J. and Sobotka. T (2012), “A Demographic Explanation for the Recent Rise in European Fertility”. *Population and Development Review*. 38 (1): 83-120.

Bryman, A. (2012), “Social Research Methods” Oxford University Press, 4rd ed., New York

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

Cameron, A.C. and Trivedi, P.K. (2009), ”Microeconometrics Using Stata”, StataCorp LP, College Station, Texas

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

Danziger, L and Neuman, S (1989), ”Intergenerational effects on fertility: Theory and evidence from Israel”. *Journal of Population Economics.* 2: 25–37

Easterlin, RA (1987), ”Birth and Fortune”. 2nd edition. Basic Books New York (first published without Chapter 10 in 1980)

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: 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

Galor, O and Weil, DN (1996), “The Gender Gap, Fertility and Growth”. *American Economic Review 86* (3):374-87

Lee, R (2003), “The Demographic Transition: Three Centuries of Fundamental change”. *Journal of Economic Perspectives.* 17 (4): 167-190

Lutz, W, Skirbekk, V and Testa, MR (2006), “The Low Fertility Trap Hypothesis: Forces that may lead to future postponement and fewer births in Europe”. *European Demographic Research Papers.*

Macunovich, D (1998), “Fertility and the Easterlin hypothesis: An assessment of the literature”. *Journal of Population Economics*. 11: 53–111

McCloskey, DN and Ziliak, ST (2009), “The cult of statistical significance”, presented at the Joint Statistical Meetings, Washington, DC, August 3rd, 2009, in a contributed session of the Section on Statistical Education. [accessed on 11 May 2015]. available at: http://www.deirdremccloskey.com/docs/jsm.pdf

McNown. R (2000), “A Cointegration Model of Age-Specific Fertility and Female Labor Supply”. Working Paper No. 00-15. *Center for Economic Analysis*. Department of Economics. University of Colorado at Boulder

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

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

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

Organisation for Economic Cooperation and Development notes (OECD) (2014). [accessed on 14 March 2015]. available at: http://www.oecd.org/els/emp/LFSNOTES_SOURCES.pdf

Pampel, FC (1993). “Relative Cohort Size and Fertility: The socio-political context of the Easterlin effect”. *American Sociological Review*. 58: 496-5I4

Potancoková, M, Sobotka, T and Philipov, D (2008), “Estimating Tempo Effect and Adjusted TFR”, Documentation to the European Demographic Datasheet 2008, Vienna Institute of Demography, Austrian Academy of Sciences

Rutten, A and Higgs, R (1984), ”Graphic Tests of Easterlin’s Hypothesis: Science or Art?” in: *Research in Population Economics,* 5: 201–212. JAI Press, Greenwich, CT

United Nations, Department of Economic and Social Affairs, Population Division (2014), ”World Population Prospects: The 2012 Revision, Methodology of the United Nations Population Estimates and Projections”, Working Paper No. ESA/P/WP. 235

Data Sources

Human Fertility Database (2015). [accessed on 13 March 2015]. available at: http://www.humanfertility.org/cgi-bin/main.php

OECDStats (2015). [accessed on 15 March 2015]. available at: http://stats.oecd.org

OECD Social Spending indicator (2015). [accessed on 17 March 2015]. doi: 10.1787/7497563b-en. available at: http://data.oecd.org/socialexp/social-spending.htm#indicator-chart

OECD Family Benefits Public Spending indicator (2015). [accessed on 17 March 2015]. doi: 10.1787/8e8b3273-en. available at: http://data.oecd.org/socialexp/family-benefits-public-spending.htm#indicator-chart

UN World Population Prospects (2012). [accessed on 14 March 2015]. available at: http://esa.un.org/wpp/Excel-Data/population.htm

7. Appendices

Appendix 1: The Low-Fertility Trap Hypothesis by Lutz, Shirbekk and Trak (2006). LFT-1 is the demographic mechanism. LFT-2 is the sociological mechanism, and LFT-3 is the economic mechanism derived from the Easterlin hypothesis.