Low fertility: unifying the theory and the demography

Peter McDonald
Demography and Sociology Program
Australian National University

Paper prepared for Session 73, Future of Fertility in Low Fertility Countries, 2002 Meeting of the Population Association of America, Atlanta, 9-11 May 2002

The market is very short-term and favors flexible, childless people averse to risk, while the private world of the family requires them to be self-sacrificing and prepared to accept risks. These two areas - which some still consider to be separate - attest to the old division of labor between men and women. The reinforcement of the male breadwinner model of the family is not the solution to the dilemma that we now face in maintaining social reproduction in combination with a free-market approach to production. What is required is no less than a new social contract asserting that children are a social good and not merely a private, optional pleasure. The new social contract also must be one based on gender equity.

Introduction: a new equilibrium or a new chaos?

Formulation of coherent theory relating to the transition of fertility from high levels (around an average of six children per woman) to low levels (around two children per woman) is one on the central challenges in the study of population. The task has brought demographers inescapably into contact with the theoretical paradigms of other disciplines, especially economics, sociology and anthropology but with associations also to population biology and ecology. There have been some useful assessments of these theories in recent years (Watkins 1986; Cleland and Wilson 1987; McDonald 1993; Hirschman 1994; McNicoll 1994; Greenhalgh 1995; Mason 1997; Casterline 2001). Central to these assessments has been the assessment of theory against empirical reality, or the demography of the transition.

Conventional demographic transition theory, in keeping with evolutionary theory, frames the historical change in fertility behavior into an equilibrium model. Prior to the transition, fertility and mortality had been in long-run equilibrium at high levels and, at the end of the transition, are again in equilibrium at low levels. At both the high and the low equilibrium, the rate of population growth is close to zero (Notestein 1945). In a broad sense, the global movement from high to low fertility can be seen as ‘convergence’ (Wilson 2001) and is consistent with a world heading towards a new demographic equilibrium. On the other hand, where fertility has fallen to replacement level, it has usually not leveled off at a new equilibrium level (around two children per woman). Instead, very low fertility rates have emerged in many post-transition societies. Present rates are so low in many countries and for many population sub-groups that continuation at these levels would eventually threaten the existence of these populations. In a shorter
time frame, they face substantial economic issues related to falling labor supplies and rapidly ageing populations (McDonald and Kippen 2001). Rather than a new equilibrium, continuation of this situation would be an example of the achievement of a new chaos following the chaos of rapid population growth in the past two centuries.

In a stable population with a Total Fertility Rate of 1.3 births per woman, the population falls at the rate of 1.5 per cent per annum. Such a population, in 100 years would fall to less than a quarter of its original size. In contrast, with a TFR of 1.9, the rate of decline in a stable population is only 0.2 per cent per annum and the population size after 100 years would be 82 per cent of its original size. Thus, it is an error to convey the impression that, in the long run of history, a TFR of 1.3 and a TFR of 1.9 is much the same thing (demographic convergence). Fertility falls from 1.9 to 1.3 through 60 per cent of women having one fewer child. Thus, a relatively small change in fertility has very major implications for population futures. Replacement migration and increases in labor force participation rates can be successful ways to avoid hyper-ageing and future falls in population and labour supply but only if fertility is in the range of about 1.6 to 2.0 births per woman. Where fertility is below about 1.6, the levels of migration required for replacement of the population become impossibly large. Impossibly large numbers of immigrants may also be required where fertility is above 1.6 births per woman if the country has a very large population size (McDonald and Kippen 2001, United Nations 2001a). The capacity for social absorption of immigrants can be a further limiting factor even when the capacity to absorb is viable from a demographic perspective.

Low fertility countries today can be divided largely into those that have fertility rates in the range of 1.7-2.0 births per woman (moderately low fertility) and those with fertility rates of 1.5 births per woman or lower (very low fertility). Around 2000, there were only a few countries with rates lying between these ranges and, in these cases, their fertility rates might be considered to be in transition from the higher to the lower category (the United Kingdom, Canada) or from the lower to the higher category (Sweden and Belgium). All countries with a fertility rate of 1.5 or lower reported to the United Nations in 1999 that they considered their fertility rate to be too low. Only one low-fertility country with a fertility rate above 1.5, France, reported that its fertility rate was too low (United Nations 2001b). More recently, France’s fertility rate has risen to 1.9 births per woman at which level the French Government may no longer consider the fertility rate to be too low. Thus, it seems that governments have recognized the demographic implications of below replacement fertility that are described in the previous paragraphs: a rate of 1.7-2.0 is acceptable; a rate below 1.5 is not. When Sweden’s fertility rate fell to 1.5, the Swedish Government instituted an enquiry (Childbearing in Focus). Thus, the important theoretical question is what are the features that distinguish post-transition countries and sub-groups with moderately low fertility rates from those with very low fertility rates. There has been a rush of empirical work in recent times, especially in European countries, that enables a closer refinement of theory relating to low fertility. The purpose of this paper is to attempt so far as possible to unify theory with this empirical work. In doing so, however, a healthy level of skepticism for absolute precision is warranted:
The standard social science model is that society works pretty much like a regression equation: the task is to find the right set of predictors, solve the equation, and discover what factors are most important in predicting social outcomes (Hirschman 1994: 226).

**Is very low fertility a temporary cross-sectional aberration?**

As will be discussed in detail below, the emergence of low fertility in OECD countries has been associated with substantial deferment of the first birth. From around 1980, the average age at first birth has risen by about 2-3 years in most countries (Lesthaeghe and Moors 2000: Table 3). This gives rise to the question as to whether current Period Total Fertility Rates (PTFR) are temporarily low because births have been postponed to a future time. When the progressive deferment of first births comes to an end, birth rates at younger ages (under age 30) will stop falling while at the same time birth rates at older ages will continue to rise, at least for a time. The effect would be that, in the future, PTFR would rise ‘of its own accord’. If the PTFR were to rise into the 1.7-2.0 range, the ‘safe’ range, then there would be no need for policy intervention to raise fertility, particularly when such policy is likely to be costly. Indeed, a temporary lowering of the PTFR may have beneficial effects in crowded countries, particularly those with an age structure that maintains a momentum for population growth. A future rise of PTFR because of past postponement implies that Cohort Total Fertility Rates (CTFR) will remain above current PTFRs. The process described here was referred to by Ryder (1964) as demographic translation (see also Keilman 1999).

The argument that very low fertility will disappear with the ending of deferral of the first birth has been made most strongly by Bongaarts and Feeney (1998). Less optimistic results have been obtained by later studies using different methods (Lesthaeghe and Willems 1999; Kohler and Philipov 2001; Kohler and Ortega 2001). van Imhoff (2001) provides a very good summary of the differences in the methods that have been used and provides his own view that it is impossible to infer cohort fertility measures from period fertility measures, simply because the cohort has not completed its experience and a range of future outcomes always remains possible. Nevertheless, he himself is more confident about some projections than about others based upon inspection of detailed data. He asserts that the best way to project cohort fertility is to do it directly using trends in cohort fertility, not indirectly from trends in PTFR. Such projections need to be based on detailed data relating to the fertility of cohorts, preferably cohort parity progression probabilities, if possible expressed in terms of duration since last birth. Inspection of these detailed series of data enables an assessment about the ending of the phenomenon of deferral of the first birth and the re-emergence of some stability. van Imhoff sees stability arising for the Netherlands enabling a more confident assessment that CTFR in the Netherlands has remained within the ‘comfort zone’, that is, above 1.7 births per woman. However, he is agnostic about the trend in CTFR for Italy on the grounds that an end to the deferral of the first birth is not yet sufficiently in view. On the other hand, it seems certain that van Imhoff would predict that the underlying CTFR in Italy is higher than the current and recent PTFR. The question is: how much higher? While van Imhoff does not draw this conclusion, an implication from his analysis is that future cross-
sectional births are better projected through the estimation of the future fertility of cohorts. Ortega (2002) has recently provided a very interesting methodology that is based upon this approach.

**The two factors in low fertility: deferral of the first birth and recuperation after a late first birth**

The discussion in the previous section indicates that CTFR can be considered as having two components: the timing and eventual probability of having a first birth and the level of childbirth subsequent to the first birth. This division has been investigated in an important paper by Lesthaeghe and Moors (2000; see also, Lesthaeghe 2001). They find that some countries have very strong levels of ‘recuperation’, that is, rises in fertility at older ages that tend to offset the falls at younger ages. For example, the Netherlands and Norway were found to have very strong recuperation while Italy and Spain had negative recuperation (rates at older ages fell as well). In eastern and central European countries, fertility in the 1990s has fallen sharply at all ages producing the lowest levels of PTFR in the world (Avdeev 2001; Philipov 2001). Kohler, Billari and Ortega (2001) argue that the fall of fertility rates in Spain and Italy to very low levels is highly associated with the increased delay of the first birth. However, at least for cohorts near to the end of their childbearing, it is evident that the factor that distinguishes societies with very low fertility from those with moderately low fertility is ‘recuperation’ rather than the effects of deferral of the first birth. In Lesthaeghe and Moors analysis, the Netherlands had the oldest age at first birth, yet it was in the moderately low fertility category because of its high level of recuperation. Coleman and Garssen (2002: 14) conclude that ‘the Netherlands may point the way for the rest of Europe’ through the age pattern of its childbearing - a high-peaked curve with a small variance and a mean around 32 years. Later in this paper, theory related to low fertility is considered in terms of:

- deferral of the first birth and childlessness, and
- recuperation, or the extent and speed of having children beyond the first.

In making this distinction, it is important to point out that recent research indicates that higher-order parity progression rates are not independent of lower-order progression rates, especially progression to the first birth (Kravdal 2002).

**The dependent variable or how to measure low fertility**

Most often, movements in fertility are reported and interpreted in terms of the trend in the Period Total Fertility Rate. This measure translates readily into the number of births occurring in a year and therefore is useful for this purpose, however, as already discussed, when the timing of births is changing the PTFR is very likely to give a misleading impression of the underlying (or cohort) course of fertility. Nevertheless, if the PTFR in one country has been 1.8 for the past 15 years and in another 1.2 for the past 15 years, as is the case between different populations in Europe, it would be safe to conclude that there is a substantive difference between their underlying levels of fertility. Some theoretical interpretations are based on such comparisons (McDonald 2000a).
Then, there is Sweden. Sweden’s PTFR has fluctuated from 1.6 in 1984, to 2.1 in 1991, to 1.5 in 1998, and is again on the rise. In contrast, the PTFR in Norway has been almost constant since 1988 at just over 1.8 births per woman. Careful studies of Swedish fertility trends indicate that the rise in the 1980s was due to a change in the timing of second births (shortening of the inter-birth interval) while the fall in the 1990s was due to a sharp rise in the age at first birth. Andersson (2002: 81) concludes that the trend in PTFR for Sweden gives a highly misleading picture of the underlying cohort fertility trend in that country and that the trend in PTFR for Norway provides a better indication of Sweden’s underlying trend.

To avoid misleading results based on crude measures such as PTFR, European research on the factors associated with low fertility utilizes more specific measures. The decomposition of fertility levels and patterns into two components, the deferral of the first birth and the extent of recuperation, has focused empirical studies on cohort parity progression rates: the chance that a woman will have a first birth by a given age, the chance that, given first birth at a particular age, she will have a second birth, and so on. Cohort parity progression rates provide reliable measures of fertility behavior and, as they are directly related to decision-making in regard to a particular child, they are also more useful in testing specific hypotheses about fertility behavior. The countries where such measures have been used tend to be countries in northern and western Europe with population registration systems that enable cohort fertility histories to be traced and, sometimes, to be linked to the characteristics of the person. Unfortunately, such data are not available for many countries and their replication through surveys requires extremely large surveys. At present, with Rebecca Kippen, I am investigating the possibility that such measures can be constructed from Australian census household records using an approach based on the own-children method.

**Heterogeneity or how interpretations of demographic trends can be misleading**

Measures available for the entire country are more reliable as dependent variables when the country is homogeneous in regard to culture and economic status. Where a country is characterized by heterogeneity, national level trends can be misleading (Vaupel and Yashin 2001). In a heterogeneous population what appears to be a change in behavior for the population as a whole can in fact be a shift in the composition of the population towards the group that displays the behavior in question. Within groups, there may be no change in behavior. Three main compositional characteristics stand out for investigation: education groups, ethnic groups and geographic groups (urban/rural). Fertility in the United States in the past 15 years has been affected by a shift in the ethnicity of the US population towards persons of Hispanic origin who have relatively high fertility. Women in the US who were born in Mexico have twice the fertility rate of non-Hispanic whites born in the United States (U.S. Bureau of the Census. The course of future US fertility is dominated by the question as to whether the fertility of this Hispanic sub-group will remain relatively high. Any trend for the country as a whole is confounded by what is happening in the ethnic sub-groups. Likewise, in Singapore, there is a wide difference in fertility rates by ethnicity (Chinese PTFR is 1.3, non-Chinese is 2.3). Fertility trends in
Singapore are therefore influenced by changes in the ethnic composition of the population and by trends in fertility for the two ethnic categories.

In relation to fertility behavior, a more universal source of heterogeneity is education level. The hypothesis could be made that below replacement fertility is simply the result of major changes in the educational composition of the female population. It is conventional to find that more highly educated women are more likely not to have a first child and, if they do, to have the first child at a much later age than women with lower levels of education. As more women shift into the higher education category, more women take on these fertility-lowering characteristics and the fertility rate for the society as a whole falls. This hypothesis is certainly worth testing, but the evidence, such as it is, suggests that there is more to low fertility than simply a shift in the education distribution of the population. Beets and Dourleijn (2001) have estimated that half of the increase in age at first birth in the Netherlands in recent decades can be attributed to the structural shift in the population towards persons with higher levels of education. The other half relates to changes of behavior within education classes. In Australia, between 1986 and 1996, the fall in fertility for cohorts was greater in absolute terms for women with low education than it was for those who had a university degree (McDonald 1998).

A counter argument to the heterogeneity argument is that people may maintain their fertility behavior when their characteristic (education, for example) changes. This would mean that the fertility behavior of higher educated women would tend to move towards that of other women as more and more women become educated. In a more extreme example of this possibility, many women in Australia had obtained their university degrees after they had had their children (McDonald 1998). The specific effects of education on deferral of the first birth and on recuperation in selected instances are considered below; here the point is simply that the analyst needs to consider the potential confounding effects of population heterogeneity on fertility outcome measures.

The utility of children: changing preferences and values

Any theory relating to fertility behavior must address the issue of why people have children. This can be expressed as the value that they place upon having a child, the utility of the child or their preferences for having a certain number of children. It is not only the preferences of women that matter but those of their partners as well. Preferences are likely to be more important when fertility is low; the choice between one and no children, or between one and two children is likely to carry more emotional weight than the choice between five and six children.

In the context of contemporary low fertility, Coleman states that, while the costs of children can be figured in dollar terms, there are no dollar benefits. Instead, the benefits or utility of children consist of dimensions of a psychological nature that are not as readily quantifiable. He refers to these benefits as ‘immanent values’ (Coleman 1998: 20). One way to think about the utility of a child in this circumstance is in terms of net benefit thresholds (the psychological benefits less the psychological costs). That is, people have some calculus of the psychological gain (utility) to them of having the next
child. If the costs of children rise or if the shape of the utility function shifts in favor of other goods, some individual benefit thresholds will be crossed and decisions will be made not to have the next child. There is also an argument that when the first child is delayed, tastes for goods other than children are enhanced shifting the utility function towards goods other than children.

We know surprisingly little about why people want to have children despite its critical importance to the determination of present-day fertility levels. However, we know that the value of children varies massively across individuals. There are some who would subject themselves to the most torturous privations in order to have a child while other would not have a child or another child under any circumstances. It also needs to be remembered that children come in highly discontinuous lumps and that the utility of having a child will vary according to the birth order of the child. Having the first child provides the benefits of the status of being a parent, of ‘being a family’, of having offspring who will carry on the line, of meeting the expectations of others, of fulfilling the narcissistic tendency to recreate oneself, of having a baby who will be fun and will grow up and love you, of fulfilling childhood dreams, of providing vicarious pleasure from the child’s success, etc. The decision to have a second child may be more related to the strength of the notion that each child should have at least one sibling, or to wanting a child of the other sex. Those who have a third child may value at least three children as a ‘real’ family, or they may be still trying for the other sex that they don’t have yet. Those who have a fourth child may simply love children, or it may have been a mistake. It is likely that the level of the net psychological benefits threshold falls as birth order rises. That is, the highest psychological threshold relates to the first child. Also, it is very likely that the level of the threshold falls as people get older. That is, all other things being equal, a woman at age 29 may feel more inclined to have a second child than a woman at age 39. Psychological costs probably rise with age or, perhaps, as argued above, increased age leads to shifts in the utility function towards other goods. Accordingly, as ages at childbearing increase, people will be less likely to have additional children.

Evidence from many studies of fertility preferences in low fertility countries suggests that achieved fertility falls short of preferences, particularly preferences that were held when women were in their early twenties. For example, van Peer (2000) provides data on desired and realized fertility for several European countries showing that preferences fall short of achievement. Similar conclusions have been drawn for European countries by van de Kaa (2001). Bongaarts (2001) also reports that when fertility is low, it is conventional to find that desired fertility is above realized fertility. Retherford, Ogawa and Sakamoto (1996) concluded that, in Japan, it has been economic and social changes that have driven fertility down and that values have lagged behind these changes. As education levels and labor force participation of Japanese women increased, preferences for children tended to remain relatively stable. While there is much debate about the meaning of different measures of fertility intentions or preferences, the message is clear that, in most low fertility countries, there appears to be a latent demand for children. This situation has provided some analysts with optimism that, despite postponement of the first birth, preferences will eventually be met (Bongaarts and Feeney 1998) and PTFR will rise.
Counter to these studies, Avdeev (2001) has argued convincingly that a strong one-child preference has emerged in the Russian Federation. In countries where the level of post first birth fertility (re recuperation) is low, while there may be a preference for two children, the level of commitment to this preference may not be high in the face of economic costs. So the issue may not simply be a stated preference, but the degree of commitment to that preference. Also, in Australia, intended family size has been found to fall as women age through their twenties and into their thirties, that is, as women confront the realities of the combination of work and family and the realities of relationship formation (McDonald 1998). Hence, post-hoc rationalization of preferences is likely. Finally, using different longitudinal surveys, McDonald and Evans (2000) and Qu and Weston (2001) have found that the preferences of Australian women tend to be realized or remain intact when the relationship that they are in at Time 1 remains intact at Time 2. The same applies to women who are not in a relationship at both Times 1 and 2. However, where relationship status changes between Time 1 and Time 2, intentions are not met and preferences change. This work emphasizes the central role played by relationship formation and stability in meeting fertility preferences.

As a final note of caution, the importance of preferences can be overstated. Referring to Southern Europeans in Australia, Santow and Bracher (1999) have made the interesting argument, largely confirmed in a recent analysis by McDonald (2002), that preferences about the number of children may be culturally weak in comparison to other aspects of family formation such as leaving home before marriage, cohabitation or ex-nuptial fertility or in comparison with expectations about economic and social mobility. As such, it is possible to imagine shifts in childbearing preferences generated by fashion or peer influence. The very wide swings in the timing of the first birth in western countries during the 20th century could not have been driven by an economic model only or by preferences that were steeped in long-standing culture. In this context, van de Kaa (2001) has speculated that, in a postmodern future, a reaction to late and low childbearing is possible, that is, people will place a higher value upon children compared to other outcomes in their lives.

The context: the interaction of globalization and local institutions and values

Varying levels of below replacement fertility can be seen as varying outcomes of the interaction of global forces coming in contact with local institutions and values. The global forces include the international flows of capital, technology, education, knowledge and values, and their accompanying institutional arrangements including changes in governance, financial institutions, labor markets, occupational structures, and employment conditions. These bring with them an associated global values structure. At the local or societal level, these global forces rub up against long-standing or traditional institutions and values. Research has shown that there are some traditional institutions and values that are very resilient to the incursion of global forces, and these are primarily the institutions that impinge upon intimate relationships, specifically, religion and family (McDonald 1994; Mayer 2001; Baizan et al 2002). Because these traditional institutions vary from place to place in substance and in strength, the outcomes from globalization
vary within each context. Accordingly, while these social conditions prevail, we can expect divergent rather than convergent outcomes for fertility.

I have argued that very low fertility (1.5 or less) in most instances is characteristic of countries or societies in which there has been considerable improvement in women’s opportunities in education and the labor force deriving from globalization and modernization, while there has been little change in women’s roles as mothers and wives as defined by the traditional institutions of family and religion (McDonald 2000a, 2000b). In such circumstances, where society and the family provide obstacles to the employment of wives and mothers in the form of unsupportive working conditions, long working hours, lack of alternative forms of child care, superior male rights and power within the family, and unequal division of household responsibilities, some women will opt not to marry, not to have children or not to have an additional child. More broadly, where women are treated as autonomous individuals in the education system and in the labor market but as inferior beings in other social institutions founded on a male-dominated family system, some women will opt to be less family-oriented than they otherwise would have been. It is in these circumstances that we can predict very low fertility as the outcome.

In association with gender inequity, as argued in the following two sections, low fertility is related to the phenomenon of delay or eschewal of family formation that has arisen as a result of the increased level of risk for young people ensuing from the deregulation of labor markets and economies.

Globalization and the new labor market

A central feature of globalization has been the emergence of a new labor market that is competitive and insecure but potentially lucrative for the most successful workers. Jobs that were once secure, lifetime positions have become insecure under labor market reforms that make it easier for employees to be fired or downgraded. Computerization has meant that many new jobs are in small, high-risk firms and contracting out of services has given rise to a broader range of short-term contract or casual jobs. Previously, high risk was associated primarily with low-level jobs. What is different now is that high risk has been extended across and up the labor market and restructuring has led to an accompanying shift of jobs to the qualified and credentialed workers. These new labor market conditions have affected young workers to a greater extent than older workers because older workers often have defined benefits as part of their employment that are not easily removed. New entrants, however, have not been afforded the same job protections. The rate at which companies come and go is also on a steeply rising curve. Most of the bankruptcies are in small companies and are largely unseen but we have also witnessed the demise of some very large corporations, a visibility that promotes the perception of risk. Globalization also means that decisions about a person’s job are often made very remotely and impersonally. On the other hand, the new labor market also offers great financial rewards to those who are successful in its terms. The deregulated labor market increases inequality and, hence, provides great incentive for young people to strive to be on the upper end of the Lorenz curve.
Investment in personal human capital is the most effective hedge for individuals against the risks of the new labor market. In the new economy, if a person's level of human capital is high, he or she can maximize returns while employed but also regain employment more quickly when the work or job ends. Young people today expect to change jobs relatively frequently and expectations are borne out by experience. This is not an argument only about high level positions and skills; the same argument applies at lower levels as well. For example, a person working in the retail industry today is expected to have sales skills, people skills, computer skills and perceptive skills that enable them to report trends observed on the shop floor to higher levels. These skills are gained increasingly through academic education, on the job training and cumulated work experience. As security is normally higher at higher levels of the firm, the employee behind the counter needs to have an eye to promotion to a more secure position. Likewise, a tradesman today relies upon knowledge of more advanced skills obtained through tertiary education and the accumulation of a reputation for good and reliable work. The accumulation of personal human capital, as argued below, conflicts with the pursuit of family formation, especially for women in societies where the level of gender equity is low.

This depiction of the new labor market and its impacts on young people is obviously starker where rates of unemployment are high for young people as in some countries of Europe especially in East and Central Europe but also in the South. Unemployment simply makes the argument about the risks of the new labor market more immediately obvious to young people and its effects upon fertility more direct.

The family and the market

Whether related or not, the new identification of workers corresponded neatly with the restructuring of economies in the 1980s in line with a philosophy that the free operation of the market is the most efficient and effective form of economic organization. In the past 20 years, regulations and restrictions have been reduced so that capital can flow easily in the direction that maximizes business efficiency and profit. The theory is that profitable businesses mean improvements in employment and wages and, hence, in economic wellbeing. The characteristics of this new economic regime are small government and low taxation, free flow of capital across international boundaries, free trade, freedom for employers and workers to determine wages and working conditions, and curtailment of government-funded social welfare. In distributional terms, the system rewards innovation and hard work and, hence, provides incentives for both. The individual worker has greater freedom to sell his or her skills to the highest bidder. Governments, both national and international, take on a new role as facilitators and regulators of the system. And employers have no interest in the family status of their workers and, accordingly, feel little or no responsibility for workers’ families.

In the 1990s, it can be argued that the free market system, on average, has produced greater levels of prosperity in most industrialized countries. However, there are questions about its distributional outcomes. While handsomely rewarding those who are successful
in its terms, the new market economy tends to be unforgiving of its casualties and laggards whether they be individuals, companies or nations. Companies and nations are penalized for less than acceptable performance through the outflow of capital at short notice. Individuals are penalized for less than acceptable performance by loss of their income source, or by stagnation in their career path. The rewards may be greater under this system than under the former system of protection, but the risks are also greater. The casualties of the system become a problem for the system only in so far as they disrupt the free operation of the market. Most advanced countries deal with the casualties of the system through their welfare systems, but these are increasingly seen as being under threat because they absorb capital that could be used more productively. Like the market itself, life for the workers has become much more of a gamble with high stakes.

The market approach deals with individuals as inputs to the system of production. Consequently, in order to protect themselves from risk, individuals must maximize their utility to the market. This means that they need to focus upon the acquisition of saleable skills, work experience and a marketable reputation. At the same time, they need to accumulate savings or wealth as a personal safety net. They also need to maintain flexibility of time and place so that they can react to opportunities as they arise. The canny player in a game that rewards market production is unwise to devote time or money to social reproduction. Social reproduction involves altruism, that is, time and money devoted to others or to the society at large. For the risk averse in a free market economy, altruism is equivalent to foolhardiness (Folbre 2001).

Family is the heart of social reproduction. It is the place where altruism abounds. There are people and politicians who believe that the public world of the market economy and the private world of the family are separate worlds: that an individual can be highly competitive, individualistic and risk averse in the market but then, just hours afterwards, be self-sacrificing, altruistic and risk accepting within the family. The only explanation I can give for this logic is that these are people who still believe in the rigid separation of the roles of men and women; that market production is a male responsibility and social reproduction is a female responsibility and that the male breadwinner model of the family is the way that people do (or should) lead their family lives. This thinking is much more likely to apply in societies that continue to hold to older models of the family that are characterized by gender inequity.

Young women today are equipped for market production at a level at least equivalent to young men and employers are very happy to employ women in the market economy. Where human capital counts, the free market will employ a skilled woman before an unskilled man, even before a man slightly less skilled than the woman. The risk-averse woman of today will ensure that she is able to support herself and, where there is a high probability of divorce, will be careful not to put her self at the risk of dependency upon a man. Couples recognize that dual employment provides a hedge against job loss for either one and banks reinforce this by providing housing mortgages on the basis of two incomes. Parents and schools encourage young women to accumulate skills that will enable them to remain attached to the labor force. As a result, there are very few young women in today’s modern economies who see their future lives in terms of finding a
husband and never thereafter being engaged in market work. Consolidation of the male breadwinner model of the family is not the solution to the dilemma that is faced in maintaining social reproduction in combination with a free market approach to economic production.

The demographic outcomes of the emergence of the new economy are longer dependence of children upon their parents, increased participation and longer durations in education, difficulty in forming relationships, delayed marriage or non-marriage, delay of childbearing and childlessness. Countries with very low fertility experience all of these outcomes to varying degrees. Similarities between countries are no longer regional. Researchers have drawn attention to the similarity of experience in recent years of Japan and Italy (Dalla Zuanna et al. 1998), and similar trends are evident in Singapore. In Japan and Italy, young people remain at home with their parents into their thirties, marriage rates are low and fertility is in the order of 1.2-1.3 births per woman. In both countries, young women are well educated and have access to well-paid jobs. They do not have to find a husband for reasons of economic security; their security is their own human capital. They have a lot to lose if they marry; the opportunity cost of marriage is high because of the nature of family and working life.

**The evidence: delay of the first birth and childlessness**

Above, I have provided a theoretical framework for the interpretation of low and very low fertility. Careful studies in Europe in particular are beginning to provide empirical evidence that supports this framework.

Baizan et al (2002) provide a comprehensive study of the lives of young people in Spain very much in the framework that I have postulated. They find that years in education have increased dramatically. From the 1950-54 birth cohort to the 1960-64 birth cohort, years in education increased from a mean of 15.4 years to 25.6 years for men and from 10.6 years to 25.4 years for women. Again between these two cohorts, the number of episodes of unemployment and the number of job changes also increased significantly. They also record an increase in the heterogeneity of experience from the older to the younger cohort, especially in employment. Finally, they observe that women's careers are becoming more similar to those of men. They conclude that all of these trends in combination with the Spanish family system and the costs of housing have led to the postponement of family and household formation. Young people, both men and women, wish to be well established in their employment before they marry and have children. Also in relation to Spain, Ahn and Mira (1998) observed that the lack of stable jobs among men is one important factor that has forced many young people to delay marriage and childbearing. Between 1987 and 1995, the proportion of employed Spanish men aged 25-39 years who held permanent work contracts fell from 55 per cent to 37 per cent. Ahn and Mira (1998: 15) concluded that the key to increased family formation in Spain lies in ‘increasing the level of confidence among young workers about their future employment prospects’.
A norm of achieving a good income situation before having children has also emerged in Sweden (Andersson 2002). Indeed, Andersson makes the interesting suggestion that where a parental leave payment is earnings-related, there is an incentive to delay the birth of the first child until a couple reaches a higher income level. de Wit and Ravanera (1998) also argue that young Canadians are inclined to wait until they are secure in work before having children but they make the additional observation that, where young people have been successful in attaining a good income and employment situation at a relatively early age, this speeds up entry to marriage and reproduction. This underlines the hypothesis of increasing heterogeneity among young people. For the Netherlands, Liefbroer using attitudinal data from a panel survey, found that the timing of the first birth is influenced by the perceived costs of having a child for one’s career opportunities but that children were also seen as reducing life’s uncertainties because of the stabilizing effect that they had on life (this in keeping with the hypothesis of Freidman, Hechter and Kanazawa 1994).

These findings sit within an interesting theoretical debate. Happel, Hill and Low (1984) presented a theoretical argument and empirical evidence using data for the United States to argue that there is a greater economic incentive for couples to postpone childbirth where women acquire high-paying jobs because of the potential loss of earnings and job skill depreciation that would ensue from time out of the labor force. Counter to this, using better measures, Kravdal (1994) found that accumulated economic and material resources have a large effect upon the timing of the first birth, whereas economic potential has little influence. Cigno and Ermisch had made the same argument for the United Kingdom but the available data were inadequate for the purpose. The reconciliation between the two arguments, as intimated originally by Happel, Hill and Low, may be the capacity to purchase childcare and other child-related needs that comes with the accumulation of wealth and the acquisition of a high income-earning husband.

As already mentioned, Beets and Dourleijn (2001) have documented the increase in durations and levels of education in the Netherlands and its impact on the timing of the first birth. Britta Hoem (2000), in explaining the fall in fertility in Sweden in the 1990s, described a remarkable shift towards education among young women in Sweden as the Swedish economy came under increased pressure. In 1989, 14 per cent of Swedish women aged 21-24 years received an educational allowance that is payable to all adult students (ages 20-50 years); in 1996, the figure was 41 per cent. At ages 25-28, the equivalent change was from 9 per cent to 22 per cent. Using municipal data, she also observed that delays of childbearing were positively correlated with regional levels of unemployment. Similar observations for Sweden have been made by Andersson and Liu (2001).

These findings suggest that, as education levels continue to rise in response to the demands of the liberalised labour market, first births will be delayed even longer. With very lengthy delays, the chance that the first birth does not occur at all increases. This becomes more the case where young people have a poor understanding of the decline in fecundity (the biological capacity to reproduce) as women age though their thirties. Beets and Dourleijn (2001) have documented the relatively poor knowledge of young people
about this issue in the Netherlands and suggest that information of declining fecundity should be included in school curricula along with family planning information, that is, they consider that young people may need conception education as well as contraception education. If this is the case in the Netherlands where information and education on sexuality is highly advanced, how much more likely is it to apply in other countries?

**Recuperation: the evidence on having second and third births**

While levels of childlessness in advanced countries may be more divergent in the future, the evidence until now suggests that the main factor in fertility difference between countries with moderately low fertility and countries with very low fertility is the extent to which childbearing continues beyond the first birth after the first birth occurs at a late age (recuperation). Recent studies in Europe have focused upon the determinants and speed of progression from the first to the second birth and from the second to the third birth. The evidence suggests that a higher education level does not lead to lower progression rates at these birth orders. Indeed, it is not unusual to find the opposite effect. For example, Kreyenfeld (2002) found a positive correlation between a woman’s education level and the transition rate to the second child for West German women. Kravdal (2001) argues that we should pay attention to the combined effects of all of the parity progression rates because of selectivity at lower progressions. He argues that educated women in Norway have a higher level of childlessness and a later age at first birth that contribute to lower fertility rates overall despite only small differences in progression rates by education at higher orders of birth. The later age at first birth of educated women means that it is somewhat artificial to examine rates of parity progression at higher ages while controlling for current age.

There seems to be strong agreement among researchers that the transitions rates from first to second birth and from second to third birth are highly related to access to resources that enable women to combine work and family. Baizan, Michielin and Billari (2002: 202) argue in respect of Spain that there is a high opportunity cost associated with childbearing because of the lack of ‘social care services’. Kravdal (1996) found that education had a less negative impact on fertility in municipalities in Norway where access to childcare was better. Ronsen (2001) concludes that the improvement of policies to support work and family in Norway has led to a reduction in fertility differences between women of different education levels. Hoem, Prskawetz and Neyer (2001) find higher rates of transition from the second to the third birth in Sweden than in Austria but find little difference in the educational levels of women in the two countries nor in the levels of individual autonomy of women. In keeping with gender equity theory (Joshi 1998; McDonald 2000a), they conclude that the difference between the two countries is brought about by public policy in relation to work and family. The opportunity cost of the third child is greater in Austria because of lack of access to resources that support the ability of women to combine work with a third child. These resources include availability of part-time work, access to affordable childcare, access to long-duration parental leave and the level of maternity leave payments. Olah (2001) draws similar conclusions on the basis of a careful study of transition from the first to the second birth, this time comparing Sweden with Hungary. Rindfuss, Morgan and Offutt (1996: 288) argue that
fertility in the United States has remained relatively high because childcare centers have become more available and acceptable. They say that ‘the preference, the need, and the ability to pay for center-based childcare is greatest among female college graduates’. This has meant that the depressing effects that increasing education of women might have had upon higher-order fertility have been mitigated. Tsuya (2000) attributes low fertility rates in Japan and South Korea to lack of support for working women both outside and inside the household.

In contrast to the strong conclusions about the effectiveness of work-family support resources in maintaining fertility rates, studies of the impacts of direct economic incentives are much less favorable (Gautier and Hatzius 1997). However, there have been claims that direct financial incentives can be effective, albeit at great cost (Lutz 1999; Milligan 2002). Direct financial subsidies assist with the direct costs of children whereas policies that enable women to combine work with family reduce opportunity costs (Ermisch 1989). Opportunity costs rise with a woman’s wage whereas direct costs of children are less responsive to rising incomes (except in so far as wealthier parents have higher discretionary expenditure on children). This means that as the wage rate rises, women will be more likely to favor the combination of work and childcare than that of staying at home and receiving a direct cost subsidy for children. Across the lifetime, it is almost certain that income earned by women when they are supported to remain attached to the labor force will be higher than any likely level of government cash transfer. The higher a woman’s income, the more likely this is to be the case. In addition, where a woman remains attached to the labor force, governments receive a greater return in terms of taxes paid and in the return on public investment in the woman’s human capital. Furthermore, research has indicated that where a woman is attached to the labor force at the time a marriage breaks down, she is less likely to make calls upon public assistance as a sole parent. All in all, as the educational level and human capital of women rises, the arguments in favor of spending limited public funds on measures to assist women to remain attached to the labor force rather than on direct financial subsidies to families with children seem overwhelming.

**Conclusion**

The review of empirical evidence undertaken in this paper suggests the following determinants of low fertility.

First and foremost, as indicated by the countries of East and Central Europe and past levels of unemployment among young people in Southern European countries, unfavorable immediate economic circumstances for young people lead to very low fertility. Globalization raises economic aspirations without necessarily providing the means to achieve those aspirations. There is evidence from Russia that a prolonged period in this situation can lower fertility aspirations leading to the emergence of a one-child family preference. One-child families are inevitable for many couples in Southern European countries and the future will tell if this leads to a change in preferences in these countries. Improvement of fertility levels in Eastern and Central Europe requires the emergence of much more favorable economic circumstances. Relative labor shortages
arising from low fertility in some of the wealthier countries of Europe could change this situation through the spread of investment into the countries of East and Central Europe. Alternatively, labor shortages in the wealthier countries may draw away the more talented workers from Eastern and Central Europe in which case these countries will have neither the investment nor the skilled workers. Very low fertility combined with substantial out-migration of young people produces a double-whammy effect on births. Which of the two alternatives applies obviously has major implications for the countries concerned. On current trends, results are likely to be mixed (Philipov 2001). Prospects in Southern European countries are much more favorable in part because of their membership of the European Union. This means that Southern European countries need to be considering the policy options for very low fertility that relate to the wealthier countries.

In the wealthier countries of Europe, Asia, North America and Oceania, very low fertility is the result of an extended delay of the first birth and, subsequently, relatively low levels of progression to the second and the third birth (recuperation). While presented as separate phenomena, there are links between the two. Most obviously, delay of the first birth means that many women are considering higher order births at a time when their physiological capacity to reproduce is falling sharply. One-child families are often the outcome. There is also a selection factor. It seems that among higher educated women, those with low psychological benefit thresholds for children are very likely not to have any children at all. Those that have a first child then become selective of women (and men) who have a stronger preference for children. Accordingly, controlling for current age, progression rates from the first to the second child and from the second to the third child tend to be as high for educated women as for those with low education levels. The caveat, controlling for current age, is important however because more educated women are much more likely to have their first birth at a later age, another feature of the linkage between successive birth-order progression rates.

There is little evidence that learned preferences are the reason for very low fertility in wealthy countries. Young people, in the main, do not graduate from their families of origin and from the education system with low fertility preferences. A preference for two children remains dominant among women and men in their early twenties. This means that very low fertility is more the product of constraint than of preference. Aside from the macroeconomic disadvantages of very low fertility (McDonald and Kippen 2001), the frustration of people’s preferences to have children does not make for a healthy society. Indeed, wealthy societies in which fertility is very low can be seen as abrogating the human rights of their citizens.

The perceived indirect or opportunity cost of having children (lost earnings) appears to be the central constraint that leads to differing fertility levels in wealthy countries. There is less evidence about the salience of the direct costs of children, although, in expensive cities, additional housing costs for those with children are likely to be relevant. It seems that people consider that, if they have confidence in their capacity to earn income and to maintain their attachment to the labor force, they will be able to meet the direct costs of at least two children. There is also evidence that couples gain this confidence in two ways:
first, by being in this situation at present (having human capital, two good jobs and some accumulated wealth) and; second, by living in a country that provides guarantees that, having had children, the couple will be able to have time to care for their children and to maintain the labor force attachment of both the father and the mother. In the absence of these circumstances and where social arrangements prevent the formation of unions other than marriage, as in Southern Europe and the wealthy countries of Asia, late or non-marriage becomes the manifestation of very low fertility. In countries where cohabitation before marriage is common, lack of confidence about economic futures and the need to pursue individual economic security puts pressure upon relationships leading to their breakdown. For example, over recent decades, cohabiting first unions in Australia have become much more likely to end in breakdown rather than in marriage (McDonald 2000c). Formation of a viable new relationship then takes time delaying the progression to family formation. Despite this, with some exceptions, those countries in which cohabitation before marriage is common tend to have moderately low levels of fertility while those where cohabitation is largely taboo tend to have very low fertility.

The policy conclusion from the above is that, in wealthy countries, incentives in the form of reduced opportunity costs of having children will be more effective than incentives in the form of cash subsidies (subsidizing direct costs). Beyond the logic of the theory and the empirical evidence at the individual couple level, this conclusion is confirmed by a comparison of wealthy countries with moderately low fertility with those with very low fertility. The conclusion is also confirmed by the failure of substantial direct financial incentives in various levels and forms to influence the fertility level of Singapore Chinese. After years of such attempts, the fertility rate of Singapore Chinese today is 1.3 births per woman. Countries with moderately low fertility are countries that have implemented major programs that support the combination of work and family. Effective policies include leave arrangements sufficient to meet income needs while a couple has two children, guaranteed return to work with an option of reduced hours, and subsidized, high quality childcare (early childhood education, in fact). Family-friendly workplaces and encouragement of higher levels of gender equity at home are also part of the package.

As argued in this paper, development of such enlightened social programs is not a natural outcome of the new market-based global economy with its deregulation of capital and labor markets and its emphasis on competition. Indeed, the opposite is the case. The new market economy despises social expenditure on the grounds that it is 'unproductive' and encourages individuals to invest in their own human capital rather than in social reproduction. At the same time, in the 1990s, the new market-based economy was effective in bringing about higher levels of economic growth and improved living standards in many countries. Some compromise is necessary. What is required is no less than a new social contract asserting that children are a social good and not merely a private, optional pleasure (Livi Bacci 2001). Such a social contract would provide confidence to young people that they will not be heavily penalized by the market if they have children. The new social contract also must be one based on gender equity.
References


Beets, Gijs and Edith Dourleijn. 2001. “Low and late fertility is expected to continue: will new population policy measures interfere?,” paper presented to the XXIVth IUSSP General Population Conference, 18-24 August 2001 in Salvador, Bahia, Brazil.


Coleman, David and Joop Garssen. 2002. “The Netherlands: paradigm or exception in Western Europe’s demography,” manuscript, Oxford University, Department of Social Policy and Social Work.


Kravdal, Oystein. 2002. “Is the previously reported increase in second- and higher-order birth rates in Norway and Sweden from the mid 1970s real or a result of inadequate estimation methods?” *Demographic Research* 6, article 9: 241-260.


