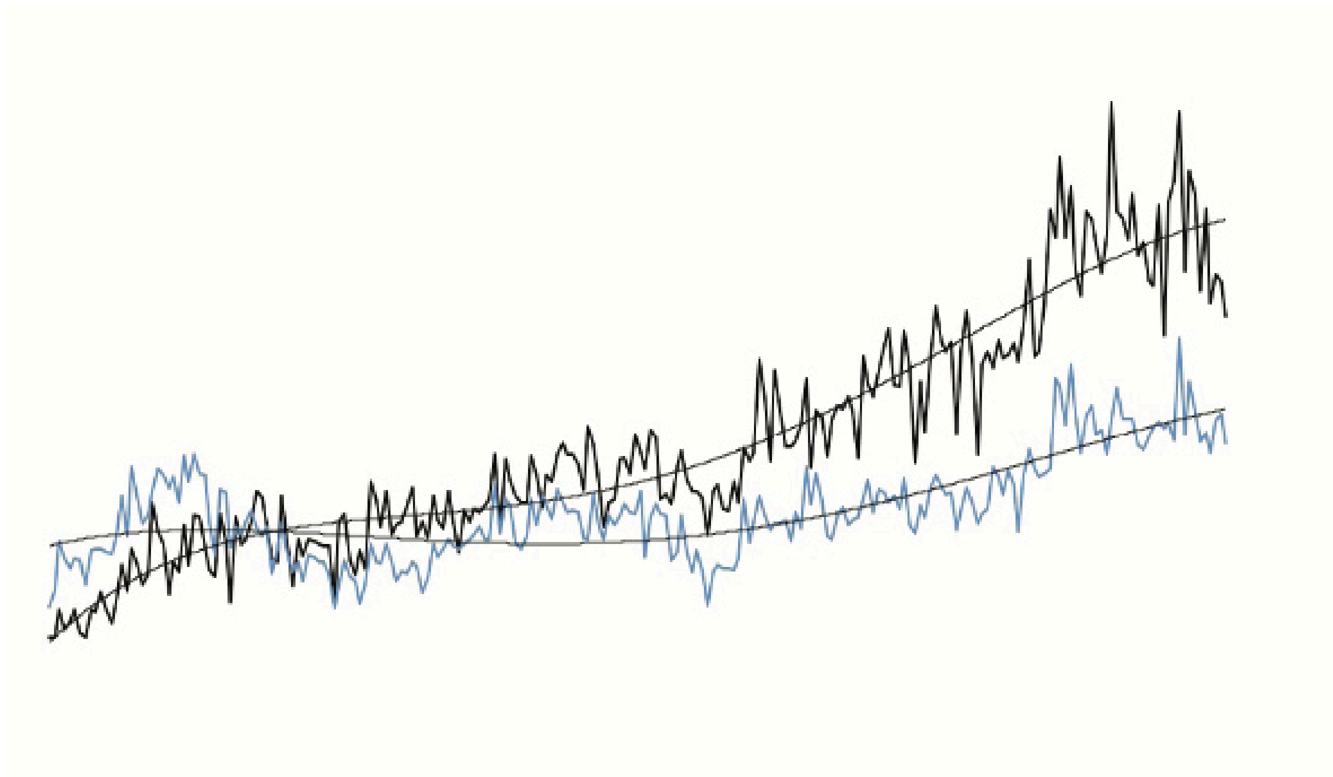
ALPHA SOURCES

AUGUST 2022



INTO THE WILDERNESS

Whether you're an evolutionary biologist, cultural sociologist or a neoclassical economist, the study of human fertility behaviour can be boiled down to an interplay between two immovable forces: the quantum and tempo effect. The first treats the fundamental question of reproduction; how many children to have, and how much resources to invest in each of them. In its simplest form, the quantum effect is the study of how much, if at all, women exert control over the quantity of offspring they produce. The extent to which they do—and almost all disciplines agree that they do in most social contexts—the analysis focuses on the conditions that determine the number of children, and how much resources that are devoted to each of them. It is an analysis of trade-offs, concentrated on the trade-off between the quantity and quality of offspring. How this balance is achieved represents one of the most crucial processes in the study of reproduction, aggregate fertility, and the demographic transition.

The tempo effect, by contrast, has a more specific meaning, at least in the context of the social sciences. It refers to the timing of the first birth, and more specifically, the observed phenomenon of birth postponement, and how this can explain the transition to low, and in some cases very low, fertility in many advanced economies, in the latter part of the 20th century. In a reproductive sense, however, it is also possible to speak of a general tempo effect, referring to the spacing between births, and the mechanisms that drive this. In a modern context, the tempo effect is important primarily because it has a lagged effect on quantum. In other words, it cuts women's reproductive career short at the bottom, with the upper ceiling determined by the menopause. This has significant consequences for the analysis of female cohorts currently in their reproductive prime. The tempo effect is important to study because it is the primary reason why fertility in one country after the other has slipped below the replacement level, and stayed there. The study of the tempo effect is implicitly linked to the question that has been vexing evolutionary biologists for two decades; are modern fertility trends fitness maximising or maladaptive?

The scope of inquiry implied by the introduction above means that a proper analysis of fertility can only be done with the benefit of a thorough understanding of four distinct disciplines; evolutionary biology and psychology, economics and sociology and their intersections. That's difficult, if not impossible, so we need to progress with precision and creativity to get a handle on the argument.

First, I will unveil a theoretical framework through which to analyse fertility, focusing on the basics of human ecology, and some of the fundamental questions that evolutionary theory is asking about modern fertility trends. Secondly, I will present the quantum effect of fertility drawing in particular on the work of Hillard Kaplan, Gary S. Becker and Oded Galor. I will finish with an attempt to explain below replacement fertility from an evolutionary perspective, before analysing the tempo effect of fertility in detail in the next next chapter.

THE SEARCH FOR A THEORY

David R. Vining's 1986 paper, *Social versus reproductive success: The central theoretical problem of human sociobiology*, highlights a key issue, which, as far as I am aware, is not yet fully resolved today. Vining (1986) controversially states in its introduction;

"The fundamental postulate of sociobiology is that individuals exploit favorable environments to increase their genetic representation in the next generation. The data on fertility differentials among contemporary humans are not convenient with this postulate."

More than a decade later Mulder (1998) revived this conundrum by suggesting that the demographic transition is a problem for evolutionary theory, and potentially a big one. Economic development since the middle of the 19th century appears to have broken two ironclad rules of evolution.

Firstly, people in modern societies tend to voluntarily reduce reproduction to levels which seem inconsistent with fitness maximisation. Secondly, the positive correlation between wealth and reproduction so clear in pre-transitional societies tend to reverse with the onset of modern, essentially post-Malthusian, economic development.

Mulder (1998) fields three broad, and somewhat contradictory, evolutionary hypotheses for sustained low fertility.

▶▶

1. Pure quantum effects – In modern labour markets, the success of individual offspring is highly dependent on, often costly, parental investment. This, in turn, implies that parents will tend to substitute quantity for quality, driving down birth rates. This idea, as we shall see, is well described in the literature via the work of Gary S. Becker and Hillard Kaplan.

2. Cultural imitation – In an environment where low fertility confers success, it is possible that (cultural) evolutionary processes operate via imitation, to create a sustained and society-wide incidence of low birth rates. This theory follows Boyd and Richerson (1975), but it has a much broader foundation in the context of the inquiry about fertility in social sciences, anthropology and the like.

3. Maladaptive behavior – This is explained by Mulder (1998) as follows; evolutionary mechanisms, either psychological or physiological, do not produce appropriate responses, from an evolutionary perspective, to rapidly changing external conditions in a modern society. As a result, maladaptive levels of fertility arise.

As we move through the gears below, the reader should remember these three explanations, as a foundation to hold on to.

EVOLUTIONARY ECOLOGY AND LIFE HISTORY THEORY

According to Voland (1998) the evolutionary ecology of human reproduction denotes:

"the application of natural selection theory to the study of reproductive strategies and decision-making."

Natural selection here can be understood specifically as the process by which species, or members of the same species, compete to extract energy and resources from the environment to most efficiently reproduce as successfully and as often as possible.

This, in turn, means accepting the Darwinian notion that humans' fer-

▶▶

tility decisions are governed by fitness-maximising behaviour, given the constraints and opportunities imposed by the ecological context. By the letter of the definition, fitness in this context can be summarised by the following passage from Wikipedia;

"it [fitness] describes individual reproductive success and is equal to the average contribution to the gene pool of the next generation that is made by individuals of the specified genotype or phenotype"

Fitness maximisation, as understood by Voland (1998), is then the behaviour of the individual, which ensures that his or her genes progress as far into the future as possible, subject to the constraints of the external environment. The conceptual framework used by evolutionary theory to describe and analyse reproduction is Life History Theory, or LHT.

This framework is concerned with the allocation of resources through an organism's life span. In the question of human fertility and fitness maximisation, the basic assumption is that the selfish gene(s) are selected so as to generate the optimal allocation of resources related to reproduction. Kaplan & Gangestad (2004) invokes a metaphor from economics, the budget constraint, when they note that biological reality, individuals operate within "finite energy budgets", giving rise to fundamental trade-offs.

Selection will favour the strategies for allocating energy that maximise fitness. According to Kaplan and Gangestad (2004), LHT should be understood as follows.

LHT provides a framework that addresses how, in the face of trade-offs, organisms allocate time and energy to tasks and traits in a way that maximizes their fitness. Optimal allocations vary across the life course and, hence, LHT generally concerns the evolutionary forces that shape the timing of life events involved in development, growth, reproduction, and aging.

Reproductive effort, in this context, is divided into three separate categories.

1. Mating - The effort put into finding a partner and successfully reproducing.
2. Parenting – The effort put into raising offspring.
3. Nepotism – Investing in relatives' reproductive effort.

From this, four basic allocation trade-offs can be identified.

The **first** is somatic versus reproductive energy allocation, which denotes the trade-off between investing in oneself—the accumulation of social and physical resources—ostensibly to be better prepared for reproduction tomorrow—and the decision to reproduce now, with the resources available Voland (1998). This trade-off between reproducing today, or waiting for more optimal conditions tomorrow is one of the most fundamental in the LHT framework for reproduction. The former is risky, potentially, due to resource scarcity, while the latter is risky because it may mean that you don't get to reproduce at all.

The **second** is the trade-off between devoting resources to your own reproduction or whether to help relatives reproduce assuming the role as "helper-at-the-nest" Voland (1998).

The **third** is the trade-off between mating and parental effort, a trade-off which differs fundamentally between sexes—as we saw here with Trivers (1972)—and is at the core of the theory of sexual selection. The divergence between the initial investment needed by male and females to produce offspring means that men have an incentive to spend a lot of time mating, and less time parenting. This is to the fact that on average they know that their offspring will be safe with their mothers. By contrast, female's reproductive strategies have evolved to offset the risk that they end up having to spend more energy parenting than their partner.

The **fourth** is the trade-off between investing in the quantity or quality of your offspring, a topic that dominates the literature on modern fertility trends.

In the case of all four trade-offs evolution operates on both the physiology and psychology to generate the optimal response given the external environment.

It is important to understand the profound difference between humans and other animals. In both cases, evolved reproductive behavior will be fine-tuned to maximise fitness across one or all trade-offs described by LHT, but the picture is infinitely more complex for humans than for most animals, for two reasons.

Firstly, this is because the human brain allows for complicated strategies and reasoning. Humans have the ability and technology to bargain with the future in the context of material resources, an ability that humans have had for millions, of years.

Secondly, humans today operate in an environment that is vastly different than a mere 200 years, let alone several thousands, years ago. Just consider the profound impact on the kind of trade-offs set out by LHT from the ability of modern contraceptive technology.

More generally, as discussed in detail below, the shift in the return on investment in the quantity and quality of offspring as well as the significant increase in the return on somatic investment for women, in modern labour markets, are two of the most profound changes. Specifically, I will focus on the first and fourth trade-offs below, as these are the most fundamental, from the point of view of modern fertility trends.

Voland (1998) opens by setting up a framework for understanding women's theoretical and practical fertility career. This, in effect, is an analysis of the trade-off between somatic and reproductive energy allocation. Three avenues of research are key in the literature;

the determinants of the timing of menarche and menopause, the determinants of the timing of ovulation, and the sperm load delivered by men during intercourse. The general idea in these areas of research is simple, but difficult to verify empirically in confined samples, let alone aggregate to society as a whole. Natural selection exerts its influence on our genes so as to optimize the timing of these key physiological events, dependent on the external environment. In the case of menarche, a naive interpretation of fitness maximisation would be that the earlier menarche the better—allowing for the longest possible reproductive cycle—but that flies in the face of two key considerations; one empirical and one theoretical.

The first is that the timing of first menarche differs significantly across populations and groups, pointing to an underlying biological mechanism, which responds to the external environment. For an evolutionary biologist or ecologist, that can only be natural selection. The second is that because reproduction is extremely energy intensive, starting as early as possible might not be optimal, especially not in an environment where a separate quantum effect is in play, lifting the relative value of the quality of offspring.

Ovulation is another process arguably governed by natural selection, though as Voland (1998) notes, "this assumption remains to be shown empirically." Voland (1998) does present evidence to suggest, however, that women's ovulation in relatively benign external environments is more sensitive to stress than for women in less benign environment. Similarly, the period of postpartum amenorrhea—the period of infertility during breastfeeding—is also found to be longer for "poorly fed mothers." The fitness maximising argument is that women with relatively low energy levels, or those subject to precarious external environments, need longer to recuperate after each birth. As such, an extended period of infertility during breastfeeding is one way in which natural selection can act over time to space births more optimally relative to the external environment.

More proximate processes also exist, most obviously via the observation recorded in many primary studies of agrarian populations that women, or couples, actively opt to space births as conditions become more adverse. This argument offers an important glimpse into the analytical methodology that (sometimes) arises from using natural selection to explain human behaviour in a modern context. Voland (1998) says:

"If changes in ovarian function can best be understood as biologically functional responses to socioecological conditions, it becomes clear that the terminological equivalence of "adaptive" and "healthy" or "maladaptive" and "pathological" is factually incorrect. (...) Ovarian "dysfunctioning" as a result of physical or mental stress (with anorexia nervosa as its most spectacular manifestation) requires clinical treatment but is also part of normal biological function under certain circumstances."

In simpler terms, if we dig deep enough, so-called dysfunctional behaviour, or traits considered adverse physiological and medical conditions, can be given an evolutionary explanation, particularly in the context of charting the trade-off between somatic and reproductive effort.

The argument that natural selection and evolution are responsible for mechanisms that govern the optimal spacing of births given the external environment—assuming no access to contraception—is uncontroversial. Natural selection has equipped the female body with protective mechanisms to prevent pregnancy in harsh external environments, and to more readily induce them in abundant environments. This is because adverse external conditions might exert pressure on the physiology to such an extent that it would be sub-optimal for the woman to become

▸▸

pregnant. The question we need to ask in a modern context is whether it is possible to detect changes in the environment—specifically linked to modernity—which drive shifts in ovulation, menarche and male sperm delivery and quality. And if we can, could we then apply such changes in an evolutionary context?

THE GENERAL QUANTUM EFFECT

One of the most important drivers of lower fertility in a post-Malthusian context is the realisation that fitness maximisation is not necessarily obtained by producing as many surviving offspring as possible, but by producing ones that are well endowed with skills, and perhaps even inheritable resources. This insight, a cornerstone in the economics analysis of fertility, can be derived from first principles of evolutionary theory. Voland (1998) notes:

"A naive interpretation of the Darwinian principle could lead to the erroneous assumption that natural selection has designed human beings to maximum fertility. This theory does not apply, however, because reproduction incurs costs. The biological process cannot favour both maximum fertility and maximum offspring fitness."

Voland (1998) offers a number of general examples in which the quantity-quality trade-off is operating to increase fertility. The most fundamental ones are the relative price of offspring, the return on investment in offspring, and the opportunity cost of contraception for women. The first of these assumes the position that children, in some instances, can yield an economic return in their own right, and that natural selection will tend to produce high fertility in such environments. The best general example is in labour-intensive economies, in which an additional child is a positive input into the family's or community's production function. Pre-industrial agrarian economies would seem to fit this bill, though other more isolated cases could exist in modern society too.

As an economist, it is tempting to analyse the number of children in such a context as a classic scale problem, in which there is an optimal number of offspring in a given family that maximises that unit's output. Analysing fertility through such a lens, however, runs counter to the point that children are not mere production inputs, at least not for humans. In most models of fertility, offspring use significantly more resources than they produce. In environments where this balance is tilted towards offspring as a production input, evolution will select for a high quantity of offspring.

The second driver of the quantum effect homes in on the return of investment in individual offspring. This, as we shall see, is fundamental to evolutionary and economic theories of modern low levels of fertility. The idea is simple. In hostile environments where the survival rate of offspring is low, random or outside parental control, natural selection will tend to favour high fertility. This is to say, in such an environment, it takes a relatively high reproductive rate to make sure even some of your genes make it into the future. It follows from this that individual offspring are afforded relatively little parental care or investment in such in environment, simply because the resources aren't available for it, or unproductive.

In environments where the return on investment in offspring is high—for example due to the availability of high-quality education—natural selection tends to produce relatively low fertility, with a high rate of investment in each child. Virtually all research on this topic holds that the transition from an environment in which the return on investment in offspring is low to one where it is high is one of the the dominant drivers of the demographic transition.

The final driver of the quantum effect is the economic opportunity costs of fertility, mainly for women, as opposed to biological opportunity costs. This is a very controversial phenomenon, especially in a framework where it is assumed that fitness maximisation is working primarily on the quantity of offspring. In environments where mothers incur no economic opportunity costs—for example if they don't work for income—fertility will tend to be higher. By contrast, in environments where women earn labour income, and accumulate assets, reproduction comes with significant economic costs, which tend to reduce fertility. This trade-off can be rationalised, to an extent, with fitness maximisation by assuming a shift in the preference for the quality over quantity of children.

That said, it still opens the door for the politically incorrect conclusion that to the extent that very low fertility in modern societies is not fitness maximising, it is because women ought to be working less, and reproducing more. As I showed in a previous chapter in discussion on sexual selection—see here—this issue sits at the heart of the fundamental difference between the investment needed by a human male and female to successfully reproduce. Biological opportunity costs of reproduction exist too, and are well described by the evolutionary literature. In environments where resources are abundant, natural selection will tend to produce higher fertility, while the opposite is the case in environments when resources are scarce. In hunter-gather and agrarian societies, where the social welfare of the individual and group is at the mercy of the external environment's ability to deliver sustenance, even slight shifts in the balance of resources can lead to an outsize impact of women's ability to carry viable offspring.

Because such societies are characterised by natural fertility, this is to say no contraception, natural selection tends to operate on the timing of births. This is to say, it regulates the balance between somatic and reproductive effort. One example is a hostile environment in which natural selection has conditioned women's physiology to increase the spacing of births, to maximise the chance of offspring survival. This is example is an important counterpoint to the naïve situation described above, in which a hostile environment produces a high fertility rate, to offset a high child mortality rate.

The remaining drivers of the quantum effect described are ad-hoc, though the idea that environments characterised by demographic competition, or the existence of a founder population and the possibility of expansion to unsettled areas, are characterised by high fertility is interesting. The most obvious example in this context is the experience of the early European settlers in the U.S., where data show that fertility was very high in selected communities, for example in New England. This driver of fertility is also potentially relevant in a modern context. It comes controversial in the context higher fertility in the immigrant population, driving a shift in the composition of the population. It's fair to say that this is as much a question of the perception as it is about actual empirical reality, though it is easy to imagine a scenario in which relatively high fertility can be fitness maximising, especially if you apply a group-selection perspective, even in a modern society. In other words, demographic competition is real, especially in an environment where the room for expansion is low, actual and perceived.

The final three drivers of the quantum effect can be summarised as follows. In environments, where children yield a positive return as a function of their ability to be helpers-at-the-nest, fertility will tend to be higher. In environments, where families have a specific gender preference, higher fertility is needed to reach the optimal mix of offspring, and finally in environments where parents need, and have the ability, to compensate for the death of a child also drives up fertility. Imagine a population hit by a virus that takes out adolescents. In such a situation, it’s plausible that fertility would rise as a response.

THE ECONOMICS STORY

In the 1950s, economists were preoccupied with one of the central questions discussed above; namely why fertility appeared to be falling in direct response to higher national income. That the demand for a good, in this case children, should be inversely correlated to rising income would make them so-called inferior goods, an odd assumption.

The work by Gary S. Becker is instrumental in driving economics towards a theory for understanding this conundrum. Becker’s fundamental insight is intuitive and strong; not only do families choose how many children to have but also how much to spend on them Becker (1960, 1973 and 1977). Once the quality of children enters into the utility function alongside quantity, a trade-off arises between the two, and the modeler can start to think about what drives the choice between them. Becker’s work is motivated by the demographic transition itself, and the shift from a Malthusian world, where fertility is explained by the age at marriage and the frequency of intercourse during companionship, to a world with contraceptive technology and social changes providing couples with more control over fertility decisions. In the language of economics, Becker laid the foundation for endogenizing families' fertility choice. In the simplest model, a rise in income will lead to an increase in the amount spent on children and thus, following a naive interpretation of fitness maximisation, an increase the quantity of children. In a modern economy, however, Becker speculated that the income elasticity of quality is higher than the income elasticity of quantity Becker (1977).

This is a fundamental insight, and in a modelling sense, it is a stroke of genius. Armed with this proposition, an increase in income drives a proportionally larger increase in the quality of offspring—often proxied by investment in education—compared to quantity. Becker's model also assumes that the price of one unit of quality is a positive function of quantity and vice versa. In other words, holding the level of quantity constant, the price of an additional unit of quality is a positive function of the number of children. Conversely, holding the level of quality constant, the price of an additional child is a positive function of the level of quality. Put differently, quantity and quality are substitutes in a modern society, a rise in income will tend to lift the latter at the expense of the former.

This assumption is violated if parents are allocating resources among offspring unevenly, of which two examples are most obvious; the first is if parents allocate more resources to the firstborn, and the second is if parents prefer one gender over the other.

In the 1980s Becker and Barro (1985, 1988 and 1989) extended this framework to formally take into account dynastic considerations. In other words, in these models the, by now ubiquitous representative consumer derives utility not just from her children, but also from her grand-children. In these models, the number of children is driven by market and subjective discount rates, and the degree of consumer altruism. Becker et. al (1990) extends the quantity/quality framework to a model in which parents with a high level of human capital have reduced fertility because it leads to higher income in the future, thereby raising the opportunity cost of the time parents spend with their children.

Further extensions of these models in 1970s and 1980s—see the overview by Iparraguirre (2018)—attempt to formalise the role of opportunity costs, or more specifically, the consequences of women's entry into the labour force, and the subsequent relative increase in return to somatic investment for women. Schultz (1973) and Willis (1973) are examples, as is Tzannatos and Symons (1980), which offers support for this hypothesis. In a sample of UK data from 1860 from 1970, the paper finds that a relative increase in women's income drove a corresponding rise in the opportunity cost of women’s time.

In other words, the increase in women's labour force participation, especially after the Second World War, created competition for women's time. Instead of using (most of) their time on reproduction, women could now spend it on somatic investment, in education, with accelerating relative returns in the labour market, especially compared to women not working.

This is one of the most important drivers of low fertility in a modern context. The question is whether this effect is strong enough to drive fertility sustainably below replacement level and if it does, is it maladaptive?

Finally, Cigno (1992) and Cigno and Werding (2007) develop the idea that high fertility in developing economies is due to the lack of financial markets that can be used for savings to insure against old age. High fertility in less-developed countries is seen as a hedge against destitution in old age, based on the assumption that children will be morally obliged to take care of their elders. By contrast, as the economy develops, financial markets will deepen, reducing the need for parents to stock up for children to make sure they are cared for in old age.

The relative income and age hypothesis, Easterlin (1966, 1978) and Macunovich (2002), deserve mention too, even if they are, strictly speaking, theoretical frameworks of their own rather than an extension of Becker's work. Easterlin's work has been instrumental for the way economists, and other social sciences, think about cohorts.

The relative income hypothesis explains the level of fertility for a given cohort to the link between a cohort's material aspiration in adolescence and their realised lifetime earnings. If aspirations are low relative to earnings, more resources will be devoted to children, raising birth rates, and vice versa, if material aspirations are high relative to realised earnings in adult life. In this framework, the baby boom after the Second World War was driven by the relatively low material aspirations of the generation growing up in the depression-stricken 1930s and the War, compared to this cohort's strong realised earnings in the post-war economic boom. By contrast, the children of the baby boom ended up with relatively high material aspirations, compared to their earnings, driving down birth rates due to the relatively higher share of resources devoted to non-reproductive effort.

Extrapolating this framework offers two hypotheses for the evolution of fertility over time, one of which offers some insight into modern trends. The first is that cohort fertility is cyclical, as a function of shifts in material aspirations—which are a function of the wealth of parents—relative to income. This idea was further developed with Easterlin's age structure model, as I explain below.

If shifts in material aspirations relative to income happen in the context of a falling trend in birth rates, however, a second interpretation is possible. Falling fertility over time could be the result of persistently higher material ambitions relative to earnings, which, in turn, could be driven by a more general economic shift in favour of the return on capital over labour.

This ties in Easterlin's framwork to the hypothesis put forward by the likes of Emmanuel Piketty on rising income and wealth inequality. If working age and reproductive-ready cohorts today are finding it more difficult to fulfil material needs such as secure employment with a wage that allows savings and the ability to own their own home, it would, in Easterlin's framework, be associated with falling, and potentially very low, fertility over time.

Easterlin (1978) adds a complementary idea to the relative income hypothesis via the importance of the relative size of cohorts. Easterlin (1978) hypothesises that the relative size of a birth cohort is a key variable determining the relationship between material aspirations and income. The larger the birth cohort, the more difficult the labour market conditions—thanks to high labour supply compared to demand—and the lower the life-time income. This could lead to baby-boom and baby-bust cycles, *within* a constant or falling trend.

Easterlin's work has stood up to empirical scrutiny, for the most part. Hill (2014) revisits Easterlin's relative income hypothesis with U.S. data and finds support for the idea. A meta-analysis by Waldorf and Byun (2005) adds to the support for the age-structure hypothesis. Finally, the work by Diane, J. Macunovich, mainly Macunovich (2002) and (2012), confirm Easterlin's work, with U.S. data.

As far as a unifying economic account of the fall in birth rates during the demographic transition Galor (2011) provides a useful overview of the key drivers of the demographic transition, from the perspective of economic analysis.

Galor (2011) starts, confusingly, by rejecting Becker’s initial quantity-quality analysis only to introduce it later, via a theory of human capital accumulation, closely related to the idea of accumulated embodied capital, championed by Hillard Kaplan, as I discuss below.

Galor’s rejection of the simple version of Becker’s theory is an exercise in killing a straw man, but it’s worthwhile running through it all the same. In its simplest form, as explained above, the quantum effect of fertility assumes that rising income is associated with a decline in fertility thanks to a higher income elasticity with respect to quality over quantity. This leads to two testable hypotheses; first that countries with higher income per capita should have begun their demographic transition relatively early. Similarly, within countries, the number of children will be inversely related to income. These two hypotheses are rejected.

The onset of the demographic transition in major European countries in the middle of the 1800s occurred simultaneously across countries with significant divergences in income per capita. This apparent rejection of Becker's quantity-quality trade is easily resolved by assuming that the shift in income elasticity with respect to the quality of children is, for the purpose of the naive model, driven by an exogenous change in the returns on human capital. As it turns out, Galor (2011) settles on exactly this explanation, eventually.

One of the most often cited catalysts for the demographic transition is that falling infant mortality–due to progress in combatting and containing infectious diseases—was instrumental in driving the decline in fertility during the demographic transition. Following Voland (1998), this would occur by reducing the number of babies needed to achieve the desired number of offspring. A simple hypothesis, which arises from this idea is that a reduction in mortality will be linked to lower fertility.

Galor rejects this hypothesis by pointing out that mortality fell for a significant period—about 140 years—before the decline in birth rates began. This, in turn, means that population growth initially rose as mortality fell, and that the population presumably was still locked in a Malthusian equilibrium. Indeed, an economy with low child mortality, high fertility but still low productivity is a prime example of a Malthusian trap as such an economy very soon will have too many mouths to feed.

The rejection of child mortality as a driver of lower fertility during the DT seems to me to be too simple a rejection of too simple a theory. For starters, it’s possible that the relationship between mortality and fertility is non-linear. Specifically, it is plausible fertility doesn’t begin to decline before mortality falls below a certain level. What’s more, it is difficult to escape the fact that falling mortality is a necessary, though not sufficient, condition for a prolonged decline in fertility. In the example above, I think it is reasonable to argue that falling child mortality itself is a byproduct of rising productivity. One could be a function of the other, indicating that the economy's ability to sustain higher fertility, for a while, is rising. The modern experience in developing economies currently undergoing their demographic transition seem to support this argument.

Interestingly, Galor (2011) does seem to believe that the increase in life expectancy in the 19th century, ostensibly driven by the same forces that helped to drive down infant mortality, contributed to drive the quantum of effect of fertility by increasing the time over which returns in each individual, both from parents and from somatic investment, can be accrued.

In the end, Galor draws on work by himself and David Weil from the beginning of the 2000s to reintroduce a version of Becker’s quantity-quality trade as the key driver of the onset of the demographic transition. The hypothesis is simple and intuitive. The acceleration in technological development during the industrial revolution, especially in the second half of the 19th century, was associated with a significant and sustained rise in demand for, and return on, the accumulations of human capital, driven by the need for labour specialisation.

From the point of view of households, it can be argued that it is exactly this shift in the relative return on human capital, which generates a shift in the income elasticity, that in turn drives the quantity-quality trade-off in offspring. In fact, it would have been odd if such a shift did not occur. The empirical evidence also seems to stack up. Several studies indicate that the rising demand for human capital in the industrial revolution was the prime driver of the decline in fertility in the latter part of the 19th century.

In England, the sharp increase in schooling from the 1850s onwards seems to be closely related to the subsequent slide in birth rates later in that century, and Klemp and Weisdorf (2010) provides authoritative evidence using data from English parishes from 1730-to-1830 to show that an exogenous increase in family size—due to the addition of another sibling—was associated with lower investment in human capital, proxied by schooling and children's literacy.

The rise in demand for human capital, due by the accelerated rate of technological development in the second phase of the industrial revolution is subject to self-reinforcing mechanisms, many of which are examples of the more general evolutionary trade-offs described by Voland (1998).

The reduction in the return on child labor is one of the main such mechanisms, which is an almost inevitable counterpoint to the rise in the return on human capital in the second part of the 19th century. Galor (2011) cites evidence to suggest that the return on child labour—wages paid to workers aged 10-to-14—fell by 50% from 1817 to 1872, and that this effect was more pronounced in relatively skilled occupations. This trend would have provided a strong incentive to substitute quantity for quality. Government policy to increase schooling and abolish child labour helped too. Evidence also indicates that the incentive to specialise in the context of expanding international trade and globalisation supported the trend in favour of a higher return on education and the quality of offspring.

The decline in the gender-wage gap is cited by Galor (2011) as another driver of the decline in fertil-

▸▸

ity in the early stages of the demographic transition. This is an interesting hypothesis given that this trend—effectively the entry of women into the labour force—usually is argued to operate relatively late in the demographic transition, mainly in the 50 years after the Second World War. That said, the data seem to hold up for the earlier parts, all the same. Galor (2011) presents evidence that the relative wages of women, and their literacy rates, rose in the second half of the 1800s, plausibly contributing to the decline in fertility in the same period.

Galor and Weil (1996) notes that the technological development during the industrial revolution increased the return of "brains over brawn", diluting the relative advantage that men have in the former. This argument hinges on the idea that rising income for women was strong enough to incentivise a rise in labour force participation, which, in turn, drove the initial decline in fertility during the early stages of the transition.

Galor (2011) is not a definitive economic analysis of the demographic transition, but it comes close. It is is part of a research program called unified growth theory. This school of thought argues that the technological development, which eventually broke the Malthusian chains was instrumental in driving down fertility by shifting the relationship between rising income and the quantity and quality of offspring. Put differently, the end of the Malthusian epoch didn't just allow the human population to grow significantly, as technology improved, and medical advances helped the human population to multiply. It also allowed, even incentivised, an increase in investment in human capital, in effect shifting the underlying positive link between higher income and higher fertility.

Guinnane (2011), which is supposed to be a guide for economists to the fertility transition broadly comes to the same conclusion as Galor (2011). It emphasises two, often, exaggerated drivers of the fall in fertility in the early stages of the demographic transition; the advent of contraceptive technology and the fall in child mortality.

The second broad group of drivers of lower fertility identified by Guinanne (2011) is linked to shifts in relative prices, inspired by Becker. These include the increase in the availability of education, the rise in the opportunity of reproduction for women, the rising costs of housing as a result of urbanisation, the change in child-labour laws, and the overall relative reduction in the return on child labour due to technological development.

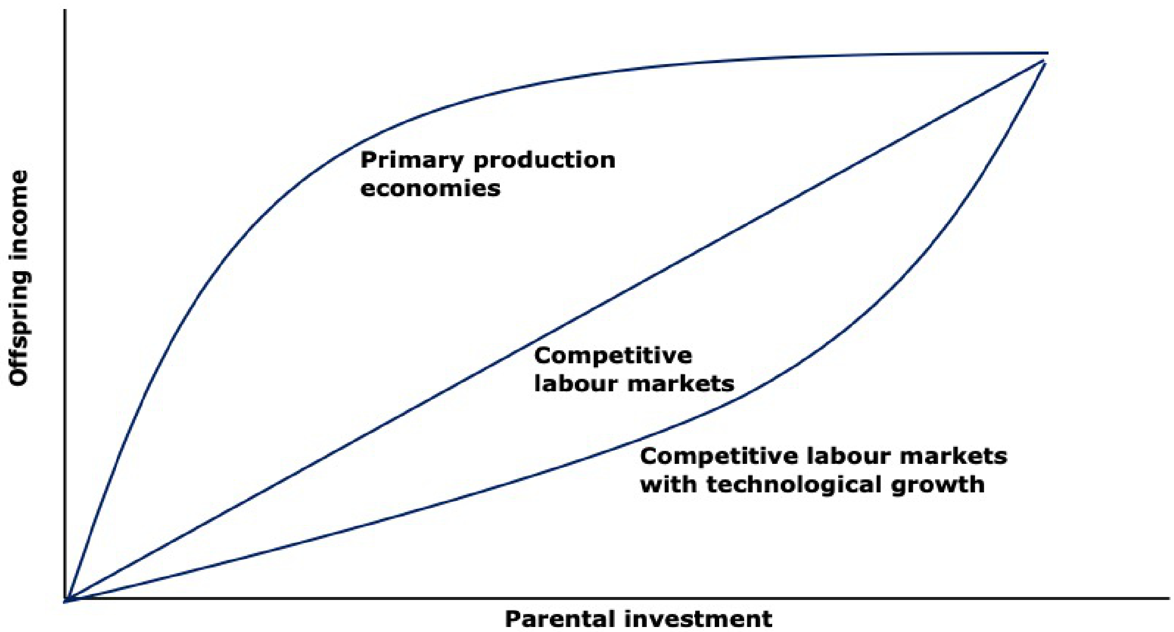
THE EVOLUTIONARY SYNTHESIS

Evolutionary theory's attempt to explain fertility trends in a modern context borrows heavily from the intuition developed by economics in the 1950s. Hillard Kaplan's 1996 article, *A theory of fertility and parental investment in traditional and modern human societies*, is one of the most comprehensive interdisciplinary theoretical treatments of the quantum effect of fertility. Kaplan's introductory observation echoes Vining (1986);

"There is mounting evidence that people in modern state societies in the developed world do not maximize fitness through their fertility decisions".

Kaplan (1996) invokes the same evolutionary conundrums which show up in Mulder (1998). Firstly, in a sample from Albuquerque, New Mexico, men are found to have less children than they ought to—given their income—and more importantly, those who have the most—apparently without incurring a fitness penalty—are parents with relatively low income.

Secondly, in industrial and modern economies, relatively well-endowed parents produce no more offspring than parents with relatively less resources, in stark contrast to overwhelming evidence from preindustrial societies of a positive relationship between income and resources and number of offspring produced.



Source: Adapted from Kaplan 1996

Kaplan's formal model to explain these observations is closely related to Becker's work, focusing on differences in which returns to parental investment on offspring's income and survival possibility—quality over quantity—differ in preindustrial and modern economies. In its simplest form, Kaplan's model suggests that investment in offspring quality, here defined as offspring income, is subject to diminishing returns in primary production economies, constant returns in a competitive economy, and increasing returns in competitive labour markets with technological growth. The chart above summarises this framework, and it is familiar picture to economists.

The x-axis denotes parents' total investment in offspring, and the y-axis is offspring income, here considered a proxy for offspring quality. For the purpose of empirical study, the first is difficult to operationalise, though investment in school and education is an obvious variable, while the second is just one out of many possible operationalisations for offspring quality.

The most obvious alternative is educational attainment, though offspring's reproductive success—the income and education of grandchildren, and great grandchildren —also come to mind. This framework lends itself to a number of empirical tests. We have to assume that any attempt to prove, or falsify, this idea will yield different results depending on the data used to bring into life parental investment and offspring quality.

The three functions above encapsulate the central regimes covered by an analysis of fertility through the lens of the quantum effect. The concave function—"primary production economies"—describes a society in which returns to parent investment in offspring quality is subject to diminishing returns. This implies that increasing investment, as a function of growing resources, is devoted to quantity instead of quality, beyond a certain point. The two remaining functions are variants of the same model. Both describe a world in which, contrary to a naive evolutionary model of fitness maximisation, reproductive investment is channeled into offspring quality instead of quantity.

The idea of varying returns to capital accumulation is a core idea in economic growth theory, especially in the transition from neoclassical growth theory to so-called endogenous growth in which capital accumulation under some circumstances exhibit constant or increasing returns to scale. In these models it is investment in human capital—often proxied by investment in education—that allows for positive scale economies.

In short, endogenous growth theory and Kaplan's quantum effect attempt to capture the same processes. Kaplan's model also incorporates the idea that returns to investment in quality accumulates over generations, an assumption Kaplan needs to account for the idea that higher wealth and income is uncorrelated with higher fertility both within and between groups. This is a strong assumption, though the implication is intuitive; it means that relatively well-endowed parents are also better at producing human capital in their offspring, and that this benefit accumulates over time. Again, we see a clear link to the ideas of dynasties developed by Becker and Barro (1985, 1988 and 1989) described above.

In summary, the fertility modes inspired by Becker and Kaplan propose two overall processes to explain the reduction in fertility during the demographic transition, both of which are identified by Galor (2011) and Guinanne (2011). First, the industrial revolution changed labour market conditions to break the link between rising income and higher fertility observed in pre-industrial societies. To re-cap, Becker proposes that the significant increase in income during the demographic transition is associated with a simultaneous rise in income elasticity with respect to quality—investment in education—over quantity. Drawing in part on Becker's work, Kaplan adds to this perspective with the idea that the transition from pre-industrial to industrial societies is associated with a shift in underlying scale economies of human capital investment. Specifically, modern labour markets are characterised by constant or even increasing returns to investment in offspring's skill, a benefit which may even accumulate over generations. In such an environment, rising income over time will tend to favour an increase in offspring's quality over quantity.

The quantum effect of fertility that arise from the work of Becker and Kaplan raises at least four issues. The first is an elephant in the room in a modern context, which I hinted at in a previous chapter—see here—on sexual selection, relying on Trivers (1972).

The relative returns to the investment in the quantity and quality of offspring is not the only trade-off that has shifted as result of economic development. The opportunity cost of reproduction has changed too, especially in a post-war context, as women entered the labour force.

This shift has driven a change in the trade-off between somatic and reproductive effort, mainly by incentivising women to earn income and assets for their own sake, and not necessarily to be better prepared for reproduction at some point in the future. Once women are able to engage in such "selfish" activities, reproduction is associated with significant economic costs, driving down fertility. Arguably, this interpretation relies on a somewhat naive evolutionary analysis, but it still lurks ominously in the modern discourse all the same, especially in the context of sustained below-replacement fertility. To the extent that such low fertility is considered sub-optimal or maladaptive, it's difficult to escape the conclusion that economic development has created perverse incentives for women to expend too much somatic, compared to reproductive, effort.

The second question concerns the empirical evidence, or lack thereof, of the quantum effect. More specifically, the quantum effect is mainly a way to explain observed behaviour. Kaplan notes:

"This analysis is a form of reverse optimality. We already know much about the basic empirical patterning of modern fertility and parental investment behavior. The goal here is to develop an optimality framework for analyzing optimal fertility and parental investment behavior in the context of a labor market economy, and then to determine the assumptions that would have to be met for the model to predict the observed behavior. Once those conditions are specified, empirical research can be conducted to determine if they hold in modern society."

Developing a framework to explain what has happened is useful in itself, but that does not necessarily help to make predictions about the future, let alone present testable hypotheses about current behaviour. In the case of the former,

▸▸

the idea of accumulated returns to investment in quality points to sustained reductions in fertility over a very long time, but there is a lower bound problem.

"So far, I have considered only fertility reduction and not the quantitative level of fertility. I have also neglected the integer constraints on fertility and have treated fertility as if it were continuous. However, we know that minimum fertility greater than zero is one. If there were excess returns to investments in embodied capital, one might expect most people to have one child. Yet evidence suggests most people consider an only child to be undesirable and have a target fertility of two or three."

Comparing this statement to global fertility trends in the past 30-to-40 years suggests that reality has overtaken the quantum effect hypothesis. As fertility falls towards the replacement level, and below, across large swathe of developed economies, the quantum effect seems to have outlived itself in terms of acting as a comprehensive framework for explaining modern fertility. This is to say, it might explain some of it, but it falls short in terms of explaining the major shifts since the 1970s and 1980s. This sets up a controversial and, as far as I can see, still-unresolved issue in evolutionary science. Either some form of shifts in trade-offs—the Becker/Kaplan quantity-quality trade-off, the change in returns for somatic and reproductive behaviour for women, or some other variant—can explain modern sub-replacement fertility or it can't. In the latter case, sub-replacement fertility is either maladaptive, or another explanation is needed.

In a general sense, a Kaplan/Becker quantum effect is inconsistent with fertility dropping too far below replacement levels, as it would be running into the time-old adage of not putting all your eggs in one basket. Specifically, there is a lower bound of one child per family because, for obvious reasons, you cannot increase invest-

▸▸

ment in the quality of something that doesn't exist. The issue, from the point of view of the individual family, is that investment a lot in a few children, if not one child, raises the risk that your investment is all for nothing in the case of injury or accident. Even if we allow for the possibility that a Becker/Kaplan quantum effect can drive fertility below replacement levels—in societies where child and adolescent mortality are low—the further birth rates fall below two, the more inclined researchers will be to ask whether some other force is in play, for good reason.

By contrast, it seems more likely that a shift in the relative returns on women's somatic investment could drive fertility below replacement levels. This is because having even one child is costly for a woman, in relative terms, who could otherwise devote those resources to her own development, most likely her career and earnings in the labour market.

For evolutionary science, below replacement fertility raises a more fundamental question about the interplay between the fertility decisions of individuals and these decisions' importance for the whole. The argument that the genes are always in control is controversial in the context of the obvious feedback between individual behaviour and the external environment. This is a particularly relevant observation in the context of economic analysis, where the external environment is subject to its own emergent properties that are often not easily reconciled with individual behaviour or even the behaviour of the otherwise ubiquitous representative agent.

In this context, the idea of selfish genes operating in the background to always and everywhere reconcile the behaviour of the individual and the group, in a general sense runs into trouble. Or doest it.

Evolutionary theory will not easily admit to this. Richard Dawkin's famous idea of the selfish gene remains a parthenon of the discipline, and it won't be easily discarded. According to Mr. Dawkins' work, evolution is driven by genes, and not, for example a separate, and explicit, group or lineage selection. I think that this is much more controversial delineation than Mr. Dawkins think.

The whole is rarely a simple sum of the parts, especially not in the world of social and economic analysis, nor I'd argue in a world where evolution operates on extended phenotypes and memes. It seems to me obvious that emergent group and society-level properties, that would not otherwise have been predicted from pinning down the behaviour of selfish genes, can and do occur.

The central question for an analysis of fertility is; if observed behaviour deviates from an evolutionary optimal path of fertility maximisation, does one go looking for an evolutionary explanation at all costs or is the "misbehaviour" a result of a temporary deviation from a pre-stated optimal path. If it is the latter, what are the drivers of such maladaptation, and how long can we expect it to last?

As far as I can see from the literature, this question continues to haunt evolutionary science to this day, posing a unique challenge for economists trying to incorporate an evolutionary perspective into their work. Economists are interested in the here and now, as well as looking ahead 20-to-30 years, for the purpose of charting the future shifts in age structure and their economic implications.

To the extent that such a period is characterised by a deviation from a fitness-maximising behaviour, and explicit stand is needed on the conflict between traits and phenomena that are temporary deviations from evolutionary optimal behaviour, maladaptations, or perfectly logical results from natural selection, if you merely look closely enough.

Voland (1998) is on Richard Dawkins' side. He notes that the selfish genes do not care about the greater good. Referring to Lee (1994), Voland (1998) notes that;

"Homeostasis and self-regulation are typical concepts in use in population history, and whenever dramatic changes are observed in demographic patterns, such as in connection with the "demographic transition," self-regulatory mechanisms are suspected of having gone off course"

Voland (1998) is sceptical about the existence of such a mechanism, for two reasons.

Firstly, Voland (1998) points out that the search for homeostatic demographic self-regulation has so far proven unsuccessful, implying that we shouldn't be looking at all.

Secondly, using the selfish gene as the unit of analysis precludes motives such as the "preservation of the species", or the "greater good". Voland (1998) concludes;

"There is neither a proximate mechanism nor an ultimate cause for population self-regulation."

Demeney (1997) draws the same conclusion. In Jones et al. (1997) he says;

"Clearly, recorded demographic experience during the past three decades lends little support to the notion that a TFR [total fertility rate] of approximately 2.1 represents a plausible temporary, let alone sustained, resting for fertility for fertility trends—a point demarcating the end of a secular process of fertility transition."

Maybe we don't have to worry too much about this. Whatever the long-run equilibrium of birth rates is, it is possible that this number has to be calculated as an average over such a long period that it will be irrelevant for most economic analysis. In any case, the evidence to date offers plenty of puzzling and interesting data for economists, chiefly related to the effects of rapid population ageing.

▶▶

Still, a population with below replacement level fertility eventually will breed itself out of existence, or at least into irrelevance. This is difficult to reconcile with the behaviour of selfish genes. More problematically, the unit of analysis—fertility–is a discrete variable in which the probability of zero increases sharply as the average falls. It is fair to say I think that a fertility of zero is not consistent with fitness maximising behaviour of selfish genes, no matter the degree to which one or more trade-offs operate to reduce fertility.

So, is there no definitive answer?

Kaplan et al. (2002), A*n Evolutionary Approach to Below Replacement Fertility*, tries to offer one. It attempts to offer an evolutionary explanation for why a seemingly maladaptive trait such as sub-replacement fertility is an enduring characteristic of modernity in one country after the other.

Kaplan et al. (2002) rely on the framework and concepts discussed above, but re-arrange them in a slightly different way, so it is worthwhile presenting the argument in full. The hypothesis is that below replacement fertility in a modern context arises with in interaction of fundamental evolutionary traits—this is to say processes which dates back to the separation of humans from apes—and two key environmental shifts in the modern economy; the emergence of extra-somatic wealth and rising returns to investment in embodied capital in oneself and one's children.

The former needs spelling out in comparison to the point made above that the opportunity cost of women's investment in reproduction has shifted in a modern context thanks to rising returns to somatic investment. Somatic investment here is understood as resources devoted to the woman herself not necessarily in order to reproduce tomorrow. Kaplan et al. (2002) defines *extra*-somatic wealth as resources that can be stored outside the body such as cattle, grain, money, property and other assets.

Embodied capital offers a high-resolution perspective on somatic wealth. It includes everything from the strength of the immune system and muscle tissue to the knowledge and skills acquired to survive. Crucially, humans invest in their own embodied capital and that of their children.

The argument made by Kaplan et al. (2002) is a version of a generic story often told by the biological sciences. Economic development in the last 200-to-300 years, have altered human conditions far more quickly than the timespan through which evolution acts. In other words, humans today operate in environment to which our brains and physiology are not fully adapted. If true, it is easy to see why maladaptive traits might emerge.

Humans, according to Kaplan et al. (2002), are characterised by five distinctive traits that separates us from other mammals; a large brain, a long life span, a long period of juvenile dependence, reproductive support from older post-reproduction family members, and male reproductive support through the provision for the female and offspring. These four traits evolved, according to Kaplan et al. (2002), thanks to a dietary shift towards foods of higher quality, in large packages, which are relatively difficult to acquire.

I am not qualified to say whether this hypothesis is true, but it doesn't have to be, to drive forward the argument.

Kaplan et al. (2002) echoes Voland (1998) in arguing that reproductive trade-offs are universal across time among humans. Pre-modern individuals, especially women, also had to allocate resources according to the trade off between reproducing now or later, the latter associated with investment in somatic embodied capital. Equally, they had to choose the optimal combination between the quantity and quality of offspring.

▸▸

Equipped with this, Kaplan et al. (2002) is able to take the leap into modernity and low fertility with the two assumptions made above; the emergence of extra-somatic wealth and the increasing returns to investment in embodied capital.

Can these two factors explain sub-replacement fertility?

As far as extra-somatic wealth is concerned, I have my doubts. Extra-somatic wealth defined by Kaplan et al. (2002) existed long before humans broke their Malthusian chains. Kim Sterelny's 2022 Aeon article, *How equality slipped away*, for example, shows how the foundations of the kind of extra-somatic wealth described by Kaplan et al. (2002) was laid almost 10,000 years as humans made the transition from hunter-gatherers to farming. Sterelny effectively argues that the emergency, in hunter-gathering communities, of complex hierarchies—mainly clans in which elites with information control emerged—provided the bridge to a more stationary, and effectively, protectionist lifestyle. Sterelny says that:

Farming and storage make inequality possible, perhaps even likely, because they tend to undermine sharing norms, establish property rights and the coercion of labour, amplify intercommunal violence, and lead to increases in social scale.

We can quibble about the importance of "farming" and "storage" as unifying concepts in this context, but it seems to me a profound insight that the development of technologies which allow humans to bargain with the future, via the accumulation of wealth and increased importance of descendants and dynasties, is a key foundation for the emergence of modern civilisation.

It's possible that such technologies changed in a way as to promote sub-replacement fertility, though they alone can't account for this phenomenon. Sophisticated extra-somatic wealth existed the onset of the demographic transition, and that they had existed for a long time.

Kaplan et al. (2002) settle on investment in embodied capital as the principal route through which fertility might fall below the replacement level. In doing so, the paper responds to the conundrum posed in Kaplan (1996). It is suggested that rising returns to investment in education across the population as a whole and the rising share of women's labour force participation are two principal reasons for fertility falling below 2.0 child per women. Specifically, it is argued that the two reinforce each other.

This intuition stretches back to Trivers (1972). In a framework where women, due to relatively large gametes, incur a higher cost of reproduction than men, a rise in women's opportunity cost of spending resources on reproduction—as other opportunities arise, mainly in the labour market—can lead to profound shifts in the number of offsprings produced.

At least three important modern interaction effects serve to reinforce this process; the advent of birth control technology, labour saving technology for household work—freeing up women's time in relative terms—and a delay in family formation for both men and women, as a consequence of the increase in investment in one's own embodied capital and due to the cost of choosing the wrong partner, especially for women.

The third of these points is crucial. Kaplan et al. (2002) say;

Thus, the model proposes that the below-replacement fertility is primarily a phenomenon of the most educated sector of our society and that it is primarily due to delay rather than a reduction in target fertility to below replacement fertility.

In other words, the reason why postponement can drive fertility below desired levels is because the delay eventually runs into the biological reality that women find it steadily harder to conceive as they approach menopause.

CONCLUSION

This chapter has covered a lot of ground, but in the end, it has tried to answer a relatively simple question. What are the drivers of falling fertility in the latter parts of the demographic transition as economic development begins?

For evolutionary theory, this is almost existential question. Vining (1986) first raised the issue, and the field is still debating it to this day. In a modern context, the link between resource abundance and fertility reverses. In the latter stages of the demographic transition, higher wealth and income tend to the associated with a reduction in fertility, and in some cases even a fall to sub-replacement fertility. Is such seemingly non-fitness maximising behaviour maladaptive, or can it be reconciled by rummaging through long enough for explanations in the evolutionary toolbox?

In economics, the genesis of the study of fertility in the 1950s similar. Becker's seminal work begins by exploring an economic explanation for why income and fertility seems to be inversely correlated in a modern context.

As it turns out, evolutionary theory and economics settle on a similar explanation in the end, summarised elegantly by the work of Hillard Kaplan, specifically Kaplan (1996) and Kaplan et al. (2002). In a modern context, income and wealth are used to increase the quality and offspring, at the expense of quantity. The income elasticity of quality is greater than the income elasticity of quantity.

This trade-off between the quantity and quality of offspring is one of many trade-offs of that determine the reproductive pattern of humans. Another such trade-off is the one faced by women in terms of devoting resources to reproduction now or to investment in her own skills and physical health. In a pre-modern context, this is a

▸▸

simple trade-off between using available resources to reproduce today, versus accumulating more resources to do it tomorrow.

Once women enter the labour force on an equal footing with men, however, this relatively simple picture is shattered, because women now are able to accumulate wealth and income. This increase in women's opportunity costs of reproduction accelerates as women become equally, and in some circumstances more, skilled than men. This, in turn, creates a strong incentive to postpone having the first child, and to have fewer altogether.

The literature points to a strong interaction effect between the generic quantum effect and the shift in women's reproductive opportunity costs. The better the education of the parent, the higher the relative return on each unit of income invested in the child's quality—most often education—compared to increasing the number of children, for a given level of quality. In short; not only are women incentivised to invest more in their own education for selfish reasons, but also in order to create the best offspring.

Reconciling this with the question posed by Vining (1986) seems simple at first. It is perfectly possible that exchanging quantity for quality is fitness-maximising. But it is equally questionable whether such a process driving fertility sustainably below replacement levels is consistent with fitness maximisation. Kaplan et al. (2002) suggest that it might be. Crucially, the fall in fertility to below replacement levels is seen, by Kaplan et al (2002) as partly a result of birth postponement, a topic I turn to next.

This space has intentionally been left blank

Find the landing page for this piece, and an overview of the project as a whole, here.