How effective are family-planning programs at improving the lives of women? Some perspectives from a vast literature

Shareen Joshi

February, 2011

Abstract
In the past 50 years, family-planning programs have been heavily promoted across the developing world. A vast academic literature now tests both the intellectual rationale for these programs, as well as their impact on a wide range of demographic and economic outcomes. In recent years, the availability of new methods and new datasets from the developing world has intensified the academic research on these issues even though the support for family-programs themselves has diminished. This paper examines the economic and demographic literature on family planning programs and summarizes evidence of their impact on fertility as well as additional outcomes such as child mortality, investments in children’s human capital, the economic status of households and the macro-impacts on communities. The goal is to provide policy-makers with an understanding of the strengths, limitations and points of agreement that emerge from this vast literature.
1. Introduction

In the past 50 years, family planning (FP) programs have been heavily promoted across the developing world as a means to reduce fertility rates and promote economic development. The central assumption behind such programs is that the decline in birth rates during the early stages of demographic transition can promote economic growth, reduce environmental pressures, reduce dependency ratios and strengthen a society’s ability to invest in health and education (Coale, Hoover, and Press 1958). At the micro-level, it has been assumed that a decline in fertility would relieve women of the burden of repeated child-bearing and free up opportunities for them to increase schooling and participate in the labor-force.

A significant literature – shaped by economists and demographers – now tests these assumptions (Kelley and McGreevey 1994; Kelley 1995). Much of the literature however, remains either theoretical or focused on macro-correlations between variables such as fertility or population growth and indicators of development such as GDP growth or female education. The causal impact of declining fertility and/or the impacts of FP programs on fertility have proved to be difficult to find. One of the main challenges faced by researchers is that fertility decline is affected by a wide range of variables, including socio-economic variables such as income, education (particularly female education) and female employment. Changes in these variables can affect the demand for FP, the structure of the programs, and their ultimate impact. There is also the issue of policy itself. FP programs are rarely rolled out randomly. Placement of programs in areas with distinct characteristics made it difficult to identify the precise policy driver of any observed change in behavior.

In recent years, the research has been enriched by the availability of new methods and new datasets from the developing world. This includes cross-sectional surveys such as the Demographic and Health Surveys (DHS), panel datasets such as the Family Life Surveys, and the use of random assignment evaluation methods that study causal relationships under careful scientific experimental structures. This paper examines this literature and summarizes evidence of the impact of FP programs on fertility as well as additional outcomes such as child mortality, investments in children’s human capital, the economic status of households and the macro-impacts on communities. We define an FP program as any organized effort to encourage couples to limit their family size, and space their births by using contraceptive information and services. This includes legislative, regulatory, and programmatic efforts to supply contraceptives to a population as well as efforts to reduce the demand for children and/or increase the demand for contraception through information and/or social marketing campaigns.

The paper is organized as follows: Section 2 provides a brief history of FP programs in the post WWII era and argues that FP programs have declined in priority after the ICPD conference in Cairo in 1994. Sections 3 and 4 provide an overview of two strands of the literature on FP programs: non-experimental studies that use cross-sectional or panel data to evaluate large-scale FP programs in states, countries or regions; and experimental studies that analyze random or pseudo-random pilot projects. Section 5 examines the literature on the cost-effectiveness of FP programs. Section 6 provides some perspectives that are likely to interest policy-makers.
2. Family planning programs: A short history

Family planning (FP) programs emerged after World War II. The world’s first major program was established in India in 1951 and was soon after followed by Pakistan, the Republic of Korea, and China. By 1975, about 74 developing countries had established them (Seltzer 2002; Cleland et al. 2006). Most programs fall into three general groups: (a) those that specifically aimed to curtail population growth through explicit policies such as promoting contraception and/or establishing incentives to have fewer children; (b) those that did not aim to curtail population growth, but promoted FP for other purposes; and (c) those with no explicit population policies but allowed outside donors to run programs that were mostly small in scale (Nortman and Hofstatter 1980; Nortman 1985).

The first group was dominated by Asian countries, mainly East Asia and some South Asian countries (Mauldin, Berelson, and Sykes 1978; Lapham and Mauldin 1985; Mauldin and Ross 1991). In China and Vietnam for example, the governments formally announced in the early 1960s that couples should have no more than two or three children and began a wide-range of interventions that either directly or indirectly contributed towards this goal. Many governments provided citizens with incentives to meet these targets. In China, couples with more than two children (or one child in some parts of China) were subjected to fines and penalties, though the enforcement of this program varies significantly (Short and Fengying 1998; Attane 2002). Permanent methods of contraception (mainly sterilization) were often provided for free. In Korea, Indonesia and Thailand, FP programs focused heavily on the expansion of usage of IUDs and other temporary methods in addition to permanent methods. In South Asia, the programs were less strong than in East Asia but large in scope. India for example, established a vast network of clinics that were to provide contraceptive services. In the 1960s, this was followed by a public health–based outreach program which emphasized education and awareness particularly in rural areas (Harkavy and Roy 1997). A common feature of almost all programs in this group was that they were generally led, funded and managed by domestic governments, and involved a broad range of ministries and mass organizations that focus on educating, promoting, and encouraging couples to use FP methods. Another common feature of programs in this group is that they were typically one component of broader development policies that aimed to increase access to health-care, education and industrialization.

---

1 These authors have developed quantitative measures of family planning program strength, or "effort" that are based on the number and quality of institutions that are involved with family planning programs. The measures of effort came from the belief that strong family planning programs must possess some essential features: (a) It should offer a full range of contraceptive methods and deliver them through several delivery systems, particularly in rural areas; (b) It should have a corps of full-time fieldworkers and educated the public about contraception; (c) Prominent leaders should issue frequent statements favoring the use of contraceptives; (d) The program should have a full-time director, placed well up in the government structure, and various ministries and private agencies should provide technical, logistical and financial assistance. More will be said about these criteria, and the studies that support them later in this paper.

2 In Vietnam, Bryant (1998: 246) writes that right before fertility declined, thousands of health workers were given basic training and sent to villages to promote use of mosquito nets, distribute locally made drugs, deliver babies, administer vaccinations, and carry out other standard primary health care functions.
The second group of countries was dominated by Latin America. Until about 1960, governments in this region remained strongly pro-natalist in their ideals (Mundigo 1996). This changed due to a concern about the high incidence of unsafe abortions in the mid-1960s. Abortions performed in unsanitary conditions by unqualified personnel were believed to contribute to maternal mortality and also resulted in large public expenditures as women with abortion-related complications sought care en masse from public hospitals (Mundigo 1996). To minimize disagreement with the Catholic Church however, FP programs in Latin America began as small private initiatives that were largely funded by international donors and NGOs. In most countries, particularly Brazil and Peru, these programs were ultimately incorporated into national public health programs. By the 1980s, countries in this group generally had broader goals than simply reducing fertility and/or the practice of unsafe abortion. They generally aimed at improving maternal and child health through greater birth spacing, access to pre- and post-natal care. Some Asian programs also fall into this category. Bangladesh is particularly noteworthy. Its national program, launched in 1976, aimed to provide women with a wide range of contraceptive methods through home-visits by a network of locally recruited female-health care workers. Sterilization was included in the package of options and in the first few years of the program, compensation was offered to those who chose the procedure (Cleland and Mauldin 1991). Yet the program remained largely voluntary and focused on maternal and child-health more broadly. A similar effort is seen in Iran, which launched its program in 1989. Free contraceptives were distributed through a network of village health workers, who also advised women on a broad range of maternal and child health issues.

The third group of countries was almost entirely dominated by Sub-Saharan Africa. Some countries did establish programs early on. Kenya and Ghana for example, established FP programs in the late 1960s. Tanzania established a FP program in 1970. Senegal established an urban FP program in 1976 and a rural program in 1979. Much of Francophone Africa however, remained largely untouched by the wave of interest in FP programs throughout this period. A 1920 French law that banned advertising and distribution of contraceptives continued to prevail. Across most of Africa, issues of population growth remained sensitive and highly politicized throughout the post-war period. Nigeria for example, adopted a national population council to study the issue of population growth but did not adopt any national policies to lower fertility (Caldwell and Caldwell, 1983). This was at least in part because census data that formed the basis of such decisions was regarded as too controversial. The results of the 1962 and 1973 census were actually nullified due to dispute and controversy over accuracy of the size of minority groups.

Even when they were adopted, African FP programs differed significantly from their counterparts in other countries in several key ways. First, the focus was almost entirely on temporary methods, since permanent methods were regarded as culturally unacceptable (J. C Caldwell and P. Caldwell 1987; J. C Caldwell and P. Caldwell 1988). The establishment of robust supply chains for temporary contraceptives

---

3 Only five Latin American countries fell in the first group – Mexico, Colombia, the Dominican Republic, El Salvador and Guatemala. Mexico is the largest among these. In 1974, access to family planning was declared as a constitutional right for all couples. In 1977, a national coordinating body was establish to expand the supply of contraception and a demographic target of population growth of no more than 2.5 percent per year by 1982 was declared. A wide variety of methods, including oral contraceptives as well as permanent sterilizations, were offered and the contraceptive prevalence rate doubled within a span of less than five years (Rodriguez-Barocio et al., 1980).
however, proved to be very challenging in the African context. The health-care infrastructure in this region was weak and burdened with a high demand for curative services. Moreover, most rural women resided far away from clinics or health centers (Caldwell and Caldwell 1992). Stringent eligibility criteria also made it difficult to reach women. In many cases, a woman’s access to contraception required the written consent of husband, proof of marital status or age, blood tests (for oral contraceptives), frequent follow-up visits, and non-evidence based requirements that she be menstruating at the time that she starts using certain methods such as IUDs or hormone-based systems (Campbell et al., 1996). Cultural preferences for high fertility often made women unwilling to be seen attending these clinics. Moreover, since consumers did not receive adequate information about contraceptives, side-effects were often misinterpreted and rumors were propagated. These factors combined to cause discouragement and discontinuation in the long-run (Campbell et al., 1996). Across Asia and Latin America, the spread of primary healthcare services, rapid increases in female schooling, the processes of socio-economic development and the use of marketing campaigns to promote awareness of FP programs may have alleviated some of these problems.

A second distinctive feature of African programs is that they were supported by a large number of international donors who rarely coordinated their actions with national governments or even between themselves. Since the weakness of domestic health infrastructure ruled out the establishment of “vertical” programs that packaged FP with primary health services, donors preferred to fund standalone programs that they could establish, manage and monitor themselves (Seltzer, 2002; Robinson and Ross, 2007; Mayhew, Walt, Lush and Cleland, 2005). The programs thus often remained small-scale. The goals used to evaluate the programs were often short-term in keeping with the demands of short budget-cycles. This approach stands in stark contrast to Asian and Latin American programs that were typically run by Ministries of Health and were backed by long-term budget commitments.

**Donor retreat**

International interest in FP programs lost momentum in the early 1980s. The intellectual shift behind this is often referred to as “revisionist thinking” and refers to a retreat from Malthusian fears about the crippling effect of population growth on economic growth as well as the concern with the adequacy of supplies of food and natural resources (Kelley 1995; Kelley 2001). A wide range of factors fueled revisionist thought: the rapid pace of fertility decline in Asia, the success of the green revolution, the lack of convincing academic evidence for a negative relationship between population growth and economic growth, etc. Economists emphasized that the long-run impact of population growth in economic development may not necessarily be negative. On the contrary, investments in human capital and innovation in growing populations can even have positive effects on growth and development outcomes (Simon and Lincoln 1977; Boserup 1981). Critics of FP programs used this literature to make the case that many FP programs in the developing world had been conceptualized and implemented with a false sense of urgency after World War II, without sufficient internal debate, deliberation and consensus (Kelley 1995; Kelley 2001).

Revisionist intellectual thought was also reinforced by the voices of NGOs in international policy. These groups highlighted examples of FP programs that had not gone well in parts of China, India, Indonesia, Mexico, Peru, etc (for a summary, see Seltzer, 2002: 62—70). In India for example, the controversial
HITTS model (Health Department operated, incentive-based, target-oriented, time-bound, and sterilization-focused) was so unpopular that it contributed to the collapse of the Indian government in 1977 and prompted India to launch a significant critique of FP programs in international policy circles (Harkavy and Roy, 2007). Feminists argued that women in particular, paid a high price for population policies, for they had often been viewed as passive “targets” who needed to become “acceptors” of contraception (Dixon-Mueller, 1993). They demanded that policies recognize women as key agents in the process of reproduction and must empower them – through education, information and access to health services (that include but are not limited to FP) – to have establish control over their bodies.

The impact of the rights-based approach was visible at the Vienna Conference on Human Rights in 1993 where there was an explicit recognition of the importance of reproductive rights and the need for national and international development policies to be built around these rights. The biggest shift however, occurred in Cairo at the International Conference on Population and Development in 1994. The definition of reproductive rights was took center-stage and included not only issues of reproductive decision-making, but sexual health and female empowerment more generally. Delegates explicitly called for dropping demographic and FP program targets in favor of a broader policy agenda that included a range of reproductive and sexual health measures. FP thus became embedded into a broader set of policy-goals. In the words of the UNFPA,

“[The ICPD Programme of Action] places human rights and well-being explicitly at the centre of all population and sustainable development activities. The Programme of Action moves discussion beyond population numbers and demographic targets: its premise is that development objectives -- including early stabilization of population growth -- can be achieved only by basing policies and programmes on the human rights, the needs and aspirations of individual women and men. Human-centred development -- in the sense of investing in people generally, and particularly in health, education and building equity and equality between the sexes -- is seen as a firm basis for sustained economic growth and sustainable development (UNFPA, 1995:9).”

The representatives of 179 governments agreed on the need to ensure universal access to reproductive health services by the year 2015. These governments also agreed to increase spending on population

---

4 According to Harkavy and Roy (2007), the government’s Department of Family Planning estimated that more than 20 million births were averted between 1956 and 1975. Calculations based on the number of births averted concluded that the annual birthrate fell from about 42 live births per 1,000 population in 1960–61 to about 38 in 1970–71 and about 35 in 1974–75.

5 The critique of FP programs also came from health advocates who argued that despite the strides made in the safety of technologies such as oral pills and injectibles, safety issues remain. They did not agree that the benefits of choice outweighed the risks and argued that contraceptive safety needed greater attention (Seltzer, 2002). Other health advocates argued that FP had absorbed far too much policy attention and development assistance, neglecting other critically important issues.

6 This includes the right not to be alienated from their sexual or reproductive capacity and bodily integrity through coerced sex or marriage, denial of access to birth control, sterilization without informed consent, freedom from unsafe contraceptive methods, from unwanted pregnancies or coerced child bearing, from unwanted medical attention.
and related programs. The needed resources were estimated at $17 billion a year by 2000, climbing to nearly $22 billion by 2015.7

This agenda was however met with resistance by several groups. First, there were those who challenged the coupling of a gender ideology with issues of reproductive health. In Jordan for examples, elites felt that the proposals regarding reproductive health were acceptable, but proposals that aimed to reduce gender-based violence, deliver sex-education to adolescents, spread information about STIs and promote gender equality conflicted with existing cultural norms and could not be implemented quickly (Luke and Watkins, 2002; Seltzer, 2002).

A second challenge to the Cairo agenda came from religious groups that felt the expanded definitions of reproductive health and reproductive choice tacitly included abortion and more controversial methods of fertility reduction. Even though the Cairo agenda was carefully worded to not support abortion in any circumstances, many people believed that the two issues were too deeply related to be separated in practice (Seltzer, 2002). In 2001, with support from the Vatican, the United States publicly opposed abortion, once again implemented the “gag rule” and thereby withdrew association with all organizations that offer women abortion services as a part of their general effort to expand reproductive choices for women.9,10

A final challenge came from the sheer breadth and language of the agenda itself. Some have argued that the focus on sexual health and reproductive rights was so broad that it simply failed to gain traction in parliaments and chambers of government across the developed world (Glasier et al. 2006; Fathalla et al. 2006). While Cairo advocates emphasized the importance of rights, donors were most interested in arguments that demonstrate a clear economic return on investment (Fathalla et al. 2006). The loss of focus also led to a fragmentation of academic and policy research. Many turned their attention to new competing priorities, such as HIV (Glasier et al. 2006; Fathalla et al. 2006, Blanc and Tsui, 2005).

A visible sign of just how divisive the Cairo agenda was comes from the UN’s Millennium Development Goals, agreed to by nearly all nations in 2000. The only goals that were related to reproductive health were the reductions in maternal and child mortality. Reproductive choices and reproductive rights were completely sidelined. In fact, the Cairo goal of universal access to reproductive health services was possibly the only goal that had been agreed to through a series of global conferences that did not make the final list of eight Millennium Development Goals (United Nations, 2000). In 2007, these goals were modified to include “universal access to reproductive health” by 2015. Progress was to be measured by

---

7 [www.unfpa.org](http://www.unfpa.org)
8 The Program of Action stated that “in no case should abortion be promoted as a method of family planning,” and elsewhere that “in [such] circumstances in which abortion is not against the law, such abortion should be [made] safe.” This was intended to be a compromise between those who opposed abortion on all grounds and those governments and NGOs who permitted abortion in varying degrees.
9 This was one of President George W. Bush’s first acts in office in January, 2001.
10 A cap of $15 million was placed for foreign NGOs and multilateral organizations who could not certify that they will not support any abortion-related activities, even if they use their own funds for these activities (Population Action International, 2000).
four indicators: the contraceptive prevalence rate, the adolescent birth rate, antenatal care coverage, and the unmet need for FP (United Nations 2007; UNFPA 2011)

Another sign of the damage from Cairo is seen in the international HIV policies. In the late 1990s, policymakers in the United States and indeed much of the world, were compelled to focus on the challenge of HIV. Rather than building services into FP programs however, donors chose to establish entirely new programs. One of the biggest examples is the establishment of President’s Emergency Plan for AIDS Relief (PEPFAR) under US President George W. Bush in 2003. This was one of the largest efforts in history to address the challenge of one disease. The only relationship between this program and FP was its attempt to encourage abstinence as a form of prevention of HIV. Some have argued that the focus on HIV and AIDS simply replaced the Cairo agenda, when they should have in fact simply reinforced and complemented it (Blanc and A. O Tsui 2005).

In summary, the global interest in FP programs has swung from extreme interest after World War II to disinterest at the turn of the century. The weakness of such programs is most pronounced in Sub-Saharan Africa. The rate of contraceptive prevalence remains only 26%, less than half of the world average, despite significant investments in treating sexually transmitted diseases such as HIV (WDI 2010). At the current time however, there appears to be a renewed interest in the role of FP, particularly in approaches that are broad-based, female-focused, voluntary and respectful of basic human rights. In the section ahead, we review the literature on the effectiveness and impact of FP programs with the goal of demonstrating that such programs can have impact on not just fertility but a variety of other aspects of women’s well-being.

3. What do we know about the impact of FP programs? Perspectives from the non-experimental approach

FP programs in the 1960s, 1970s and 1980s were mostly established in the absence of scientific evidence or agreement about program “best-practices”. This is mainly because detailed time-series data on economic as well as demographic variables was scarce at both the micro- and macro-level. Over time however, data was gathered and researchers began to test some of the fundamental assumptions underlying FP programs. The first wave of studies used a non-experimental approach, i.e. they evaluated the impact of FP programs involving using cross-sectional or panel data from a country, region, or set of regions to test the hypothesis that FP programs impacted contraceptive use or fertility. As more data became available, and FP programs were rolled out, these studies were updated and expanded. The research continues to evolve today, even though the interest in FP programs has declined among policy-makers.

Most of the research in this area faces two key challenges. The first is the challenge of measurement. Given that FP programs are heterogeneous in goals, quality of services, delivery systems and implementation strategies, researchers must construct a measure of program strength and not rely

---

11 The following conditions were imposed in the United States Leadership against HIV/AIDS, Tuberculosis, and Malaria Act of 2003: 20% of funds were to be spent on prevention, 15% to be spent on palliative care, and starting in 2006, at least 55% were to be spent on treatment, at least 10% be spent on orphans and vulnerable children, and at least 33% of appropriated prevention funds be spent on abstinence-until-marriage programs.
simply on simple indicators of program presence. There is also the challenge of measuring the level of socio-economic development in a region where a program is introduced.

A second challenge faced by researchers concerns the issue omitted variables. Many aspects of socio-economic development and cultural context (particularly women’s roles and opportunities) may affect the placement of FP programs, the success of FP programs, as well as the level of fertility and pace of fertility decline. Excluding a consideration of these variables induces significant “omitted” variable bias in empirical specifications. Controlling for these factors can be challenging in small samples. In the 1970s and 1980s, many strides were made in addressing the challenge of measurement. In the past twenty years, the focus has shifted to addressing the challenge of omitted variables. This is a theme we will address later on in this paper, when we discuss the use of panel data and the use of experimental programs to study the effectiveness of FP programs. The strides made by those using cross-sectional data are summarized below.

3.1 Cross-Country Studies
Cross-country studies of FP program impacts typically test the hypothesis that a FP program has a positive effect on contraceptive prevalence rates, and/or a negative impact on fertility rates or crude birth rates. They typically acknowledge that levels of socio-economic development can also impact fertility and attempt to control for these variables. An early study along these lines was conducted by King (1974) in a report for the World Bank. The analysis of 19 countries, estimates the association between the output of FP programs (as measured by program user rates) on the one side, and socioeconomic and program input variables, on the other. The variables included in the analysis are per capita gross national product (GNP), female secondary school enrollment, death rates, proportion of population in urban areas, newspaper circulation and population density. Program input variables include the numbers of service points, physicians and other personnel, and the amount of funds expended or allocated for the FP program. The study finds that FP service points alone account for 62 percent of the total variance in program user rates. The number of service points is associated with socioeconomic variables, but the degree of dependence is not very high. Interaction effects among all the variables account for roughly a third of the total variance. With respect to variance in total user rates, socioeconomic variables appear to have greater explanatory power than program input variables, although the difference is marginal. The study concludes that on the whole, both social change and FP programs play a positive role in promoting increased contraceptive practice and a decline in fertility.

In a series of papers beginning in 1972, Lapham and Mauldin proposed that programs be evaluated based on quantitative estimates of inputs and levels of “effort”. Lapham and Mauldin (1972) consider include indicators in four categories: (1) Policy and stage-setting activities, (2) Service and service-related activities; (3) Record-keeping and evaluation; and (4) Availability and accessibility of fertility-control supplies and services. Countries are classified as having levels of program effort that are strong, moderate, weak or very-weak and levels of fertility decline are assessed accordingly. In their first paper, Lapham and Mauldin propose 15 indicators and apply it to 20 countries to demonstrate that the relationship between program “outputs” and program “inputs” can be quite complicated. In subsequent work, 15 additional measures of program effort were constructed for more than 100 countries and used to evaluate the impact of FP programs (Lapham and Mauldin 1985; Bongaarts, Mauldin, and Phillips
Socio-economic variables are also included in most analyses of the impact of FP programs. These variables typically include adult literacy, children’s school enrolment, life-expectancy at birth, infant mortality rate, male and female labor force participation in industry, GNP per-capita, the degree of urbanization, and penetration of communications technologies such as televisions and radios. The relationship between the index of socio-economic development, the FP program in a country and the decline in fertility are estimated using cross-country data. This permits the authors to explore whether the effect of a strong FP program is greater in advanced social settings and smaller in less advanced social settings.

An interesting example of this approach is the analysis by Bongaarts, Mauldin, and Phillips (1990). These authors construct a “development index” based on 4 variables: GNP per-capita, the infant mortality rate, gross enrollment rate, and a composite measure of communication (number of TVs, radios and cars per capita). Countries were classified as having levels of socio-economic development that are “High”, “Upper-Middle”, “Lower-Middle” or “Low”. Results suggest that significant declines in fertility (3.1 children per woman) have occurred in countries that ranked high along the measures of program effort and socio-economic development. Moreover, there have been almost no declines in fertility at the bottom of these rankings. Key results from Bongaarts et al. (1990) are displayed in Table 2. Similar results for earlier time-periods are reported by Lapham and Mauldin (1985) and Mauldin and Berelson (1986). All these analyses argue that the level of socio-economic development has a substantial relationship to fertility decline, but also that FP programs have a significant, independent effect over and above the effect of socioeconomic factors.

In recent years, some strides have been made in studying the impact of the quality of FP programs on observable outcomes. Jensen (1996) for example, uses Indonesian data to show that women supplied by public clinics are more likely to use contraception than their counterparts who receive services from the private sector. Some aspects of service delivery, rather than client differences, accounted for the success. Feyisetan and Ainsworth (1996) use data from Nigeria to study the relationship between contraceptive use and FP services. They find that several structural factors did affect contraceptive use: both longer pharmaceutical operating hours and greater method choice increased usage, and provider fees did not constrain demand. Mwabu, Ainsworth, and Nyamete (1993) use Kenyan data to show that higher prices may reduce demand for medical services. Hotchkiss (1998) also examines price effects in Cebu, Philippines. By combining household and facility data, Hotchkiss asks whether the interaction of price with quality creates a positive relationship between price and health service demand. His results suggest that higher prices reduce health care utilization by poorer households, but otherwise have little effect.

A central weakness of this branch of the literature is that it focuses on the relationship between service provision and contraceptive use. It is largely silent on long-term implications of the programs or their impact on fertility levels. More importantly, the issue of omitted variable bias remained unresolved. Many argued that FP programs are in fact endogenous to levels of socio-economic development and other unobserved factors that affect both the effectiveness of the programs and the demand for children. This issue was highlighted in an important paper by Schultz (1994), who combined cross-country cross-sectional data for three years (1972, 1982 and 1988) to investigate the impact of the
strength of FP programs, in addition to education, religion, occupational structure, infant mortality, and increased GNP on fertility rates. He finds that the impact of FP programs depends heavily on the statistical assumptions. If infant mortality is included as a dependent variable, the effect of program effort weakens, presumably because FP programs are largely incorporated into maternal and child-health programs in many developing countries, and so part of the program effect is displayed through the infant mortality variable. He argues that FP programs are likely to be correlated with changes in the demand for children and so it is plausible that a positive program effect may be a proxy for underlying demand, and thus be entirely endogenous.

One of the strongest proponents of the importance of socio-economic variables in driving demand for fertility is (Pritchett 1994a), who examines the role of socio-economic factors in determining desired family size using a database of Demographic and Health Surveys. Pritchett argues that nearly 90 percent of fertility decline worldwide was in fact related to “desired” family size and not the effectiveness of FP programs. He illustrated that where families face higher costs of raising children, and lower risks of child mortality, the demand for children, as measured by “desired fertility” falls. Pritchett argues that contraception is not important as a causal or independent determinant of fertility. Rather, contraceptive use is higher where fertility is lower precisely because desired fertility is lower, which leads to both lower fertility and higher contraceptive demand, and thus higher contraceptive use.

Pritchett’s argument was countered by Bongaarts (1995) who argued that the fertility must be decomposed into two components – wanted and unwanted. Unwanted fertility strongly related to wanted fertility (negative relationship) and can be significantly influenced by FP programs. He uses similar data as Pritchett to illustrate that unwanted fertility rises in response to a decline in desired fertility. Moreover, program-strength is indeed associated with a lowering of unwanted fertility. Pritchett’s argument was also criticized by others, who commented on the sample used, the regression specification and the omission of program characteristics (Pritchett, 1995). In subsequent work, Pritchett concedes that that a very significant strengthening of FP programs may indeed have a modest effect on fertility. His estimates imply that strengthening a FP program substantially (by 50 points out of a scale of 0-100) would reduce fertility by one birth (Pritchett 1994b). He argues that such a strengthening is highly infeasible in most contexts, and that the magnitude of the benefit is likely to be small. It is important to note however, that small demographic changes, such as a reduction in fertility by one child, can scale up into significant effects at the level of entire populations. If we accept Pritchett’s estimates as a lower bound, the rationale for establishing and strengthening FP programs remains intact.

Another response to Pritchett comes from those who argue that FP programs rarely limit themselves to supplying contraception alone. They also attempt to curb the demand for children through the dissemination of information, change preferences and build new social norms (Piotrow et al. 1997; Westoff and Bankole 1997). Much new evidence has emerged on these mechanisms in recent years. Westoff and Rodriguez (1993), examine these mechanisms in Kenya, where penetration of mass-media is quite significant and the FP program has used television, radio and print to disseminate information about FP and the benefits of smaller families. The authors develop an FP message score, and find that this is positively and significantly related to the proportion of respondents wanting no more children and
less strongly negatively related to respondents’ reported ideal number of children, after adjustment for other variables. In another study, Westoff and Bankole (1997) examine the impact of exposure to mass-media in seven African countries in the first half of the 1990s. They argue that such exposure – primarily through radios – is directly related to greater knowledge and use of contraception, intention to use contraception in the future, preferences for fewer children, and intention to stop childbearing.

In the past fifteen years, many studies have taken advantage of the availability of new data to examine the determinants of fertility decline. The MEASURE DHS (Demographic and Health Surveys) Project is particularly noteworthy in this regard. Since 1984, it has collected and disseminated nationally representative data on health and population in more than 90 developing countries. International comparisons are made easier by the fact that survey instruments vary minimally across countries and collect uniform information on fertility and total fertility rate (TFR), reproductive health, maternal health, child health, immunization and survival, HIV/AIDS; maternal mortality, child mortality, malaria, and nutrition among women and children. Angeles et al. (1998) use 14 of these national surveys conducted between 1988 and 1995 in 11 developing countries to examine the effects of community-level FP service access and quality on individual use of modern contraceptives. They attempt to correct for endogenous program placement by modeling the placement decision itself, and also examine women’s exposure to programs. They find that the net effect of the absence of universal service access independently lowers modern use from 25 percent to 16 percent. On the other hand, they find no significant effect of service quality on use, nor do they find that community-level access and quality measures strongly influence a woman's recent fertility.

Angeles et al. (2001) for example, construct multivariate models for fertility rates using DHS data or comparable cross-sectional data for several countries (Kenya, Morocco, Tanzania, Tunisia, Zimbabwe, Bolivia, Peru, Indonesia, the Philippines, China and India). The authors find that FP programs had small effects on reported “ideal” family-sizes, but have significant effects on the use of contraception. Excluding China and Zimbabwe, their estimates indicate that modern use would drop from 25.3% to 15.9% when one moves from universal access to the absence of access to FP. Contraceptive use does not however, seem to have translated into major reductions in fertility in most countries. If one averages across countries, the simulated effects of FP programs on fertility are relatively small indicating a 0.33 of a child increase in fertility over thirty years in the absence of program variables relative to universal access to FP. This conclusion however, needs to be interpreted carefully since this study assumed that program placement was exogenous, unlike in their previous work.

---

12 The authors use data from the Demographic and Health Surveys in all cases. In the case of Morocco, they use a long-term panel in which a large sample of women who were interviewed in 1992 and then re-interviewed in 1995. This allows the authors to examine causal effects of media exposure and subsequent family planning.

13 There is also evidence of the impact of the media on changing preferences from qualitative studies. Caldwell and Caldwell (1988) for example, provide strong evidence that, in a district of Mysore, India, studied in 1979-80, the family planning program affected both the desire to have smaller families and the adoption of contraception to achieve this goal.

14 The project is implemented by Macro International, Inc. and is funded by the United States Agency for International Development (USAID) with contributions from other donors such as UNICEF, UNFPA, WHO, and UNAIDS.
The final word on this debate may not yet be clear in the literature based on cross-country studies, but the literature overall does suggest that FP programs do succeed in lowering fertility, though the magnitude of the effect is often small, and both demand for contraception and the supply of it, are undoubtedly greatly impacted by prevailing socio-economic trends. One interpretation of these findings is that FP programs are perhaps most effective where broader processes of economic development – particularly female education, female employment and economic growth in general – are already under way. This suggests that FP indeed play an important role in the process of economic development.

3.2 Country and Regional Case Studies
Estimating the impact of a FP program within a single country, or part of a country, over time, has the advantage of greater uniformity of the program as well as underlying social, cultural and economic conditions. Early literature on the impact of FP programs largely assumed that programs were exogenously imposed and uniformly implemented. They used multivariate regression frameworks to estimate the effect of the program on fertility (and other outcomes) for geographic subdivisions such as villages or districts. Such analyses have been conducted for Chile, China, Colombia, Costa Rica, India, Indonesia, Korea, Pakistan, Taiwan, and Thailand, and some countries have been the subject of more than one analysis (Hermalin and Khadr 1996) and have found that the presence of an FP program has had a negative impact on levels of fertility, though in most instances, the size of the effect is small, presumably because fertility changes slowly. Some noteworthy studies are described below:

Kerala: An interesting and little-discussed example of the role of FP programs in fertility decline comes from Zachariah (1981), who studies the example of the Indian state of Kerala. Between 1966 and 1978, Kerala experienced a decline in its Crude Birth Rate from 37 per 1,000 to 26. According to Zachariah, about 40 per-cent of the decline can be attributed to the official FP program. He uses a multivariate regression model that includes FP program information as well as socioeconomic variables (caste, years of schooling and household expenditures, among others). His analysis indicates that the principal variables that seem to have determined fertility are education, caste and the availability of FP services. To ascertain more closely the relative influence of the FP program, Zachariah uses three approaches: comparison of Kerala with Sri Lanka; comparison of the characteristics of contraceptive users with those of women whose fertility has dropped; and analysis of the relationship between official FP output (represented by measures of program clients who adopt sterilization and conventional methods such as pills, condoms and IUDs) and fertility decline by district and time period. The results of the Kerala-Sri Lanka comparison are of particular interest. The two areas are similar in socioeconomic development, and both had relatively weak FP programs until the mid-1960s. They both have very similar use-prevalence rates for conventional contraceptive methods, but differ sharply in their sterilization rates (28 percent in Kerala, 9 percent in Sri Lanka), largely because of special sterilization camps set up in Kerala during the brief phase of coercive FP programs in India during the reign of Indira Gandhi. From 1975 to 1980, the total marital fertility rate in Kerala declined from 7.48 to 5.71, or by 1.77 births; Zachariah attributes just over half (52 per-cent) of this decline to the excess of sterilization rates in Kerala over those in Sri Lanka. The highest prevalence rates in Kerala are concentrated among women with the least education and the lowest incomes.
Thailand: Schultz (1992) finds significant effects of the Thai FP program on individual fertility. The Thai program featured a rather unique combination of efforts from private associations and the government. In his evaluation of these efforts, Schultz found a decline in fertility of 2.3 percent for every $0.05 spent in 1975 in the public planning program for every childbearing-aged woman in a province. The same outlay is associated with a 9—14 percent reduction in fertility if spent on private non-profit FP associations. This result should be interpreted cautiously: private associations were in operation for longer than the public program and may have had a larger estimated impact for a variety of reasons. It should not be concluded that private services are more effective than public ones. The benefits of FP programs can be even larger than such estimates suggest because their effect persists over time, and there may also be geographic spillovers from one province to surrounding provinces. Schultz estimates that the public and private programs accounted for about half the nationally recorded decline in fertility.

Bangladesh: Bangladesh’s decline in fertility has been widely discussed in academic as well as policy circles. Fertility in that country declined from 6.5 births per woman in the early 1970s to under 3 births per woman in 2004, an exceptional record considering that much of this decline occurred in the absence of broad socio-economic improvements, in an intensely religious country, with one of the lowest levels of female literacy and autonomy in the world. The national program evolved over several years in response to the research insights that emerged from experimental programs (these will be discussed later in this paper), but was largely characterized by the deployment of female health workers to the homes of married women to encourage the adoption of contraception and provide follow-up services. Also known as Family Welfare Assistants (FWAs), these workers ensured that all Bangladeshi women received a wide range of contraceptive choices, along with critical follow-up services. One of the earliest discussions of the impact of this program was in a World Bank publication entitled The Determinants of Reproductive Change in Bangladesh: Success in a Challenging Environment (Cleland 1994). The report argued that Bangladesh’s FP programs was almost entirely responsible for fertility decline. A similar argument has been made by Amin (1995) who demonstrates that key socio-economic variables – particularly infant mortality, poverty and landlessness, children’s schooling, insurance and female labor-force participation – changed very little in the early years of the fertility transition. Others have challenged these assertions and argue that socio-economic changes were indeed key ingredients of the program’s success. Caldwell et al. (1999) for example, conduct an anthropological study in the Chittagong Division in the southeast of the country. This was the region whose lower levels of contraceptive use had been identified in the 1993-94 Demographic and Health Survey (Mitra et al. 1994):

---

15 Thailand’s FP policy was uniquely influenced by private initiatives. In the early 1960s, while the official government policy was still pro-natalist, a group of concerned health officials, private organizations, and academics set in motion Thailand’s family planning agenda and established a highly successful demonstration project in Potharam. In 1965, several hospitals in Bangkok began offering family planning services through newly opened clinics and, in conjunction with the International Postpartum Program, began counseling maternity patients on their contraceptive options. The Thai government officially established an FP program in the early 1970s, with a significant information campaign in 1972, and a community-level contraceptive distribution scheme in 1974. Doctors and midwives were also trained in the provision of FP services at provincial hospitals (Rosenfield et al. 1982; Knodel, Chamratrithirong, and Debavalya 1987).

16 Data indicate that virtually all Bangladeshi women were contacted at least once by an FWA, and more than one third were reached at home every six months (Hossain and Phillips, 1996).
47) and the World Bank study (Cleland et al. 1994: 139-140). The authors argue that the region did experience major social and economic change, real and perceived, over the last two decades. Respondents felt that economic conditions had improved during their lives. Most also believed that the authority to make decisions shifted from families to couples and individuals, who presumably had stronger preferences for smaller families. It is however not disputed by any of the authors that the FP program provided valuable services at low costs, to help individuals meet their goals. This is evident in the estimates of contraceptive use: Between 1975 and 1997, the proportion of married women who had ever used contraception increased fivefold, from about 14 percent to nearly 70 percent (Mitra, Al-Sabir, Cross and Jamil, 2007).

China: Many studies show that China’s FP program has contributed to increased contraceptive prevalence and decreased fertility, though the extent of the program’s contribution is unclear due to the uneven implementation of the policy and the rapid pace of socio-economic development. The major policy initiatives in China were implemented in the early 1970s and early 1980s respectively. The first, launched in 1971, encouraged late marriage and childbearing, greater birth spacing and fertility limitation (Attane 2002). The second, launched in 1979, was the oft-termed policy of only one child per family. Fertility control became a constitutional duty in 1982 and the State Family Planning Commission (SFPC) drew up an official indicator of the maximum completed fertility called for by the FP policy (1.6 children per woman). The implementation of policy however, varies tremendously across China. Each province has enacted its own self-contained family-limitation regulations, leading to great variations in the content of regulations among provinces (Short and Fengying 1998). A complex set of incentives for compliance, and penalty for non-compliance remains in effect in most places, but is implemented unevenly. This makes the impact of the program difficult to measure. Most analyses however, attribute at least some of the massive increases in use of contraception and decline in fertility to the existence of and enforcement of these policies (Attane, 2002). The pace of fertility decline throughout the period of the policy has indeed been remarkable: recent data illustrates that the total birth rate has dropped from 2.9 before the policy to 1.94 in women over 35 and 1.73 in women under 35 (Ding and Hesketh 2006). An alternate explanation for China’s falling fertility has been the increased pace of socio-economic development. This, it is argued has lowered the demand for children and increased the incentives for compliance with the policy (Birdsall and Jamison 1983; Kaufman et al. 1989; Greenhalgh 1994 Ding and Hesketh 2006). This is particularly apparent in light of the fact that China’s neighbors have some of the lowest total fertility rates in the world: 1.38 for Japan, 1.04 for Singapore, and 0.91 for Hong Kong (World Development Indicators, 2009). Bongaarts and Greenhalgh (1985) even

17 The most stringent rules were applied in the cities where couples were encouraged to delay marriage until age 25 for women and 28 for men and to have no more than two children. The rules for those living in rural areas were more accommodating: The minimum age for marriage was set at 23 and 25, respectively, and maximum family size was set at three children. Urban and rural couples alike had to abide by a birth-spacing period of at least three to four years

18 Soon after the adoption of the one-child policy, however, popular resistance forced the government to relax its most stringent rules. From 1984 on, rural couples have been allowed a second child, subject to province-specific conditions (Attane 2002). Although ethnic minorities also are affected by birth-control policy, the rules are, by and large, less strict for them than for the general population.
argue that the policy would have been just as effective if it had been a two-child policy with a focus on later child-bearing and increased spacing between births.

While this review of country case-studies is far from complete or exhaustive, it is possible to conclude that they generally suggest that FP programs may indeed contribute to fertility decline even though the mechanisms of impact may be heavily influenced by socio-economic factors. Our conclusion is similar to that of Ahlburg and Diamond, who review such analyses of FP programs and conclude that "FP programs contribute to fertility decline, often substantially" (1996: 319).

3.3 Panel Studies
An important limitation of both cross-country analyses and country-based case-studies is that they are generally unable to address the issue of non-random program placement: governments typically place programs either where they are demanded most or alternately where they are demanded least and regarded as most necessary. Such non-random placement, when ignored, induces bias in ordinary least squares (OLS) regression coefficients that are often obtained in both cross-sectional analyses across countries or regions within a country. As pointed out by Rosenzweig and Wolpin (1986), the direction of the bias is unclear. If the government targets areas where there is the greatest demand for FP services, estimates will be upward biased, but if the government targets areas with the least demand, estimates will be downward biased. The use of panel data helps address this issue. Researchers can examine the effects of programs over time, conditioned on initial characteristics, and estimate the impact of the program regardless of these characteristics.

One of the best and earliest studies on the impact of FP programs was conducted by Schultz (1973), who studies the determinants of fertility decline in Taiwan. The Taiwanese case was interesting because fertility declined by about 20 percent in a short interval between 1964 and 1969. The declines, in percentage of the 1966 levels, were systematically related to age, ranging from 3 percent for 20-24-year-old women to 75 percent for 40-44-year-old women. To examine the role of FP programs, Schultz (1973) divides the regional observations for each regression evenly between those for which birth rates were high in 1964 and those for which they were low in 1964 and examines the changes in key social, economic and demographic variables across these regions. He observed that of the high 1964 birth-rate regions, those with the stronger birth control programs had the largest declines. Moreover, birth control inputs made more of a difference where the need and opportunity for them was greater, in the "high" rather than the "low" regions. With respect to the other explanatory variables, the partitioning seems to be of little significance.

The findings on Taiwan were further bolstered by Freedman et al. (1974) who point out that the decline in fertility in Taiwan is mainly attributable to a decline in marital fertility for women at each educational level, rather than to a significant improvement in educational levels. In 1968, Taiwanese women with no education had a higher rate of contraceptive practice than the upper educational stratum in countries that had not yet adopted official FP programs (such as the Philippines). The FP program narrowed the educational gap over the years: in 1965 the proportion who had ever practiced contraception ranged from 19 percent with no education to 60 percent with senior high school or more, whereas in 1973 the equivalent figures were 71 percent and 79 percent. These authors also illustrate
that in Taiwan, large-scale adoption of contraception has occurred while traditional familial values and practices persisted. For example, the proportion ever using contraception among women aged 35-39 is very little different whether the women hold traditional or "modern" familial attitudes: whether the parents chose the husband (69 percent) or the wife did (75 percent), or whether the wife expects to live with her married sons all her life (67 percent) or never (78 percent).

More recently, researchers have used fixed-effects models to estimate the impact of FP programs. These authors estimate changes in local birth probabilities as a function of changes in key explanatory variables, controlling for variables that stay constant over time (like religion, social norms about female mobility, etc.). This approach has been by Montgomery and Casterline (1993) to study the impact of Taiwan's FP program on its fertility decline. Montgomery and Casterline use demographic and socio-economic data from 1961—1981 from 361 townships in Taiwan to construct pooled cross-section, time-series models of diffusion dynamics. They find evidence for the diffusion of behaviors, particularly fertility, across townships in response to Taiwan's FP program. Their results suggest that FP programs may have a “multiplier” effect and that adoption of contraceptives may resemble the patterns of adoption of other types of innovations.19

Estimates from panel data of the impact of FP programs on fertility typically find small but nevertheless significant effects. Gertler and Molyneaux (1994) for example, study the case of Indonesia. They combine data from Indonesia’s contraceptive prevalence surveys, demographic and health surveys, employment and labor-force participation surveys and detailed statistics on Indonesia’s FP programs to show that fertility decline was undoubtedly accompanied by the adoption of modern contraception. The FP program however, explained only about 4-8% of fertility decline. Increased female education and wage-rates played a far more significant role. This estimate of the FP program's impact is however likely to be an underestimate if we consider its impact on changes in social norms and the diffusion of behavior documented in the case of Taiwan.

A similar estimate is found in the case of Colombia. Miller (2009) combines data from Colombia’s census, the demographic and health surveys and detailed information about the implementation of the Profamilia FP program to estimate the program’s impact in Colombian fertility decline. He finds that the availability of modern contraceptives allowed women to postpone their first birth and to have about 5% fewer children in their lifetime (about one-third of a child). These reductions explain only about 6–7% of the fertility decline in Colombia’s major population centers between 1964 and 1993, implying that other factors were more important determinants of women’s lifetime fertility.

19 The innovation diffusion theory as introduced by Rogers (1962) is the most frequently cited publications in this field. Rogers states that the cumulative number of adopters typically follows an S-shaped curve. The S-curve starts to rise slowly when the first innovators adopt to the innovation. Following that, the cumulative number of adopters rises somewhat faster due to the early adopters. The curve is at its steepest when the early majority and late majority successively adopt to the innovation. The curve increases at a slower rate when the laggards adopt slowly to the innovation.
3.4 Impacts beyond fertility
Since the decline in fertility can lead households to reallocate resources (including both income as well as time, particularly of women) to other purposes, fertility decline is often postulated to have other “cross-effects” (Becker and Lewis 1973). Such cross-effects can be difficult to identify.

In an early study using cross-sectional data, Hobcraft (1987) uses data from 41 developing countries from the World Fertility Survey to argue that births to young mothers and births following short birth-intervals are associated with elevated mortality risks, though these may be driven by confounding factors such as including length and intensity of breastfeeding, prematurity, and as yet unspecified biological, behavioral, environmental, socioeconomic, or health care effects that are known to cause large infant mortality differences between families. He speculated that greater access to FP programs would reduce these risky births and thus reduce the infant mortality rate. In a paper entitled “Does Family Planning Reduce Infant Mortality Rates?” Bongaarts (1987) goes on to argue that FP programs indeed lower the number of such “risky” births, suggesting that such programs would typically lower infant mortality. He also cautions however, that this effect may be offset by the fact that countries with higher levels of contraceptive prevalence tend to have higher proportions of births of order one and higher proportions of births after short birth intervals (of less than two years). Another study found that, even in Latin America, which has lower child mortality rates, spacing could reduce perinatal mortality by 14 percent (Conde-Agudelo and Belizan 2000). In more recent work, Rutstein (2003, 2005) uses DHS data from 17 developing countries to show that increasing birth intervals upto 36 months can reduce neonatal mortality and infant mortality. Increasing birth intervals upto 48 months can even lower the risk of child mortality. He concludes that birth spacing of three to five years alone could prevent up to 46 percent of infant mortality in developing countries.

Direct analyses of the cross-effects of FP programs face the same challenge of endogenous program placement that was described earlier (Schultz 1990). Most studies that have provided convincing evidence of such cross-effects come from the use of panel data or experimental data. In one of the best analyses to date, Rosenzweig and Schultz (1982) use data from Colombia to show that the local availability of clinics and hospital beds and FP expenditures per capita are associated with lower child mortality as well as lower fertility across women in urban areas. These reinforcing effects are generally statistically significant among women from age 15 to 49. These results have been recently corroborated by Miller (2009). As mentioned earlier, Miller calculates women’s levels of exposure to the Colombian FP program and estimates the impact of this exposure on a variety of outcomes. He finds that exposure of a woman to FP from age 15 to 44 is associated with a reduction in cumulative fertility of 10–12 percent in urban areas, improvements in her educational attainment of 0.3 years, and an increase in her formal employment, an inter-generational increase in her children’s schooling of 0.1 years, and a delay in a child’s first birth.

Similar studies exist for other countries. Rosenzweig and Wolpin (1982) for example, assess cross-program effects on fertility, child mortality and schooling in rural India, and find reinforcing program effects from FP clinics, dispensaries, hospitals, and secondary schools (see also Duraisamy and Malathy 1991). Rosenzweig and Wolpin (1986) also estimate the direct and cross-program effects of FP and health clinics on anthropometric indicators of child health and nutritional status in the Philippines. In
this study, however, the authors use a pooled cross-sections framework and use repeated rounds of the Laguna Survey instead of time-series data. Again, they find that after correcting for the fact that programs initially targeted in the regions with the lowest health outcomes, there appears to be a positive spillover effect of a FP program on children’s health (as measured by child height). They conclude that family size and child health thus appeared to be gross substitutes in the Laguna households, a sufficient condition for the presence of some barrios with a FP, but not a health, clinic.

Much of the literature on the impact of FP programs on measures of well-being other than fertility has however come from experimental studies. These will be discussed in the following section.

4 Insights from Experimental Studies

The history of FP programs includes several examples of pilot projects that tested systems of services and delivery systems in settings where fertility decline was expected to be slow. These innovative studies were often designed and implemented in collaboration with evaluators and researchers who agree that properly implemented true experimental designs provide the strongest evidence for assessing program effects. The distinguishing feature of the true experimental design is that the units of study – individuals, geographic regions or clinics – are randomly allocated to different interventions or “treatment” conditions, thus approximating a randomized control trial. The advantage of such a design is that random allocation of study units to comparison groups that did and did not receive the program or that received program variations can provide the most confident inference that an association between the program and outcome variables is causal rather than spurious, or driven by omitted variables (Heckman 1991; Duflo, Glennerster, and Kremer 2007).

In the past 30 years, many FP programs and/or distribution systems have been implemented in experimental frameworks. In a meta-analysis of 16 such pilot projects whose results were published prior to 1992, Bauman (1997) found that a FP was associated with a negative and significant impact on fertility, contraception or other outcome variables in 13 studies. His results are presented in Table 2. He reports that the average value of Pearson’s correlation coefficient, between program variables (as defined by the programs themselves) and the outcome variables, was .08 (SE = .28, P > .05). The average proportion of explained variance, obtained by averaging the R-squared values of each study, was .01. He concluded that the programs accounted for no more than 1% (r2 X 100 = 1.00%) of the variation in the outcome variables. While this estimate is small, it is possible however that the long-term impacts are considerably higher and/or that the number of births averted may still be considerably high. Moreover, estimates from meta-analysis also disguise the broader effects of such programs, particularly as they unfold over time. In the remainder of this section, we examine some particularly well-known experimental programs and illustrate the effects they have been known to have on fertility as well as a wide-range of other variables.

Taiwan: One of the earliest documented randomized experiments in the literature on FP is the pioneering study conducted in the city of Taichung in Taiwan from 1963 to 1966 (Freedman, Takeshita, and Sun 1964; Freedman, Takeshita, and others 1969). Three program components (direct mail brochures and information campaigns directed either to husbands wives) varied in intensity in a randomized treatment across 2389 neighborhoods within the city. The effectiveness of the 12 distinct
treatment packages on encouraging the adoption of modern contraceptives, such as IUDs, were compared with the neighborhoods provided with “nothing” or no experimental treatment. In six months contraceptive adoption rates increased more in the heavily treated than in the untreated neighborhoods. At the aggregate level, total fertility rates in Taichung City after a six month program intervention, declined more rapidly than in other cities – 6.4 percent compared to 3.1 percent. However, factors other than the program could explain the earlier decline in fertility in Taichung city.

The longer-term analysis of this program however, has been limited by its experimental design. First, spillovers of information about birth control methods to areas beyond the limits of the small neighborhoods and the ability of urban residents to obtain FP services from outside their residential neighborhoods may have eroded the power of the experimental design to identify behavioral differences caused by the Taichung city program. Second, a National Family Planning Program was initiated in Taiwan in late 1964, which soon spread across all 361 local administrative areas of the island, making it difficult to assess the long run consequences of the earlier Taichung program’s randomized social experiment (Schultz 1973, 1992). Finally, researchers are unable to examine the long-term impacts of the program because it did not include any follow up surveys. The ideal FP program evaluation should include follow up surveys to record the longer term consequences of the program, for a decade or longer.

Matlab, Bangladesh: The Matlab FP and child-health program is a rare example of a FP program that was implemented in the framework of a pseudo randomized control trial framework more than 30 years ago. It has produced a large literature that suggests that the delivery of contraceptive supplies and bi-weekly follow-up services by “health workers” to women in their homes resulted in a significant impact on women’s lives both in the short- and long-term (Phillips et al. 1988; Joshi and Schultz 2007; Schultz 2008).

The program was launched in the Matlab thana (sub-district) of Bangladesh by the International Center for Diarrhoeal Disease Research, Bangladesh in 1977. Of the 149 villages in the study area, 70 received new FP and reproductive health outreach services, while the remainder continued to receive only regular government health and FP programs. Community health workers visited married women in their homes once every two weeks to offer a choice of pills, condoms, foam tablets, or injectable contraceptives (depo-medroxy-progesterone acetate) and later the copper T intra-uterine device (Koenig et al. 1992). Nearby centers treated childhood diseases, maternal health problems or complicated side-effects from the use of contraceptives. Over time, additional services were added to the program, including the provision of maternal tetanus inoculation of all pregnant women, measles immunizations to all children from the age of nine months to five years, training of traditional birth

---

20 The impact of the Matlab child health and family planning programs on fertility, infant mortality and other aspects of well-being that were directly targeted by the program are found in Phillips, et al., (1988). More recently, Joshi and Schultz (2007) confirm that there was not only a short-term impact on fertility, but also a long-term impact. This more recent study also documents some indirect effects of the maternal and child health services and reports that the program impacted women’s health, their use of health inputs, wages and incomes as well as investments in children. They find significant positive effects that increase with greater exposure to the program.
attendants, oral rehydration therapy (ORT) for diarrhoea, and antenatal care (DeGraff et al. 1986; Phillips et al. 1988; Vincent Fauveau 1994).

The immediate impact of the program was to increase the contraceptive prevalence rate from 7 percent to 30 percent in the first year, and then past 50 percent in the years that followed (Phillips et al. 1988). The overall impact of the program was a 25 percent reduction in fertility (Phillips et al. 1988). Foster and Roy (1997) report that four years after program began in 1978, the birth rate in treatment area was 20% less than that in control area, and by 1990 the birth rate in treatment area was 25% below that in control area. They also found that the impact of the program on birth rates increased with the length of program exposure. Sinha (2005) and Joshi and Schultz (2007) find a 13% and 15% reduction in fertility in the treatment areas relative to the comparison areas. This corresponds to approximately one less child over a woman’s lifetime.

There is also evidence that the program’s impact went beyond just fertility decline. The program has been shown to decrease child mortality, particularly among less educated women (Muhuri and Preston 1991; Muhuri 1996). There is also strong evidence of decreases in maternal mortality. Most recently, Chowdhury et al. (2007) show that maternal mortality fell by 68% in the ICDDR,B service area and by 54% in the government service area over 30 years. The speed of mortality decline was faster after the MCH-FP program was introduced in the ICDDR,B treatment area. They too find substantial educational differentials for mortality: in their logistic regression framework featuring maternal mortality as a dependent variable, the odds-ratio for more than 8 years of schooling compared with no schooling were 0.30 and 0.09 respectively.

Joshi and Schultz (2007) find that women in treatment areas are significantly more likely to have BMIs in excess of 18 kg/m\(^2\) in the program than in the comparison areas.\(^{21}\) This can lower their risk of premature mortality. Using similar data from the Matlab area, (Menken, Duffy, and Kuhn 2003) find that women’s initial height and BMI have a “protective effect” on survival over the entire 20-year follow-up period. In particular, women whose BMI is one unit higher than average have 17 percent lower chance of dying over the sample period. The program has also been known to have a positive impact on female labor force participation, asset accumulation and savings rates. Schultz (2008) finds that women aged 25 to 54 report log monthly earnings a third higher in the program villages compared to the other villages, and the households in which women reside have proportionately more financial, agricultural, nonagricultural, and housing assets, more consumer durables and jewelry and household tube wells in the program villages (Schultz, 2008).

The decline in fertility together with improved reproductive and child health services also impacted children in the Matlab area. Sinha (2007) finds that the program increased the labor force participation for boys, but not girls. Joshi and Schultz (2007) find that it improved the schooling of boys, but not girls. Barham (2008) finds that children who were eligible to receive child health services experienced a 0.3

\(^{21}\) This indicator is consulted in public health as a measure of objective risk factors for mortality. Values below 18.5 are believed to be driven by deficits in calorie consumption, combined with physically demanding work, and poor health, diarrhea and inflammatory disease (WHO 1995, 2006).
standard deviation increase in cognitive functioning as measured by the Mini Mental State Exam when these children are 8-14 years of age.

Much has also been written about the model of health-care delivery itself. (R. Simmons et al. 1988) argue that the FP workers in Matlab went far beyond the simple provision of access to contraceptive information and supplies. They found that these women functioned as important “agents of change and that their presence “...helps to shift reproductive decision-making away from passivity, exposing women long secluded by the tradition of purdah to the modern notion of deliberate choice” (Simmons et al., 1988: 29). A second study, based on focus-group sessions of women in the Matlab area, illustrate that the workers: (1) are seen as trusted friends; (2) are said to help the women overcome the fear of contraception; (3) are reported to help the women overcome opposition of family members to FP; (4) provide access to contraception; and (5) help the women to consider making fertility decisions as "thinkable" (R. Simmons 1996).

In summary, the Matlab experiment produced not only declines in fertility but many additional improvements in the lives of the women it targeted. This has been widely noted and discussed in the literature on population because this decline in fertility was achieved in a rather impoverished agrarian society, with low female-education, low female labor-force participation and rather stringent restriction of female mobility.

Navrongo, Ghana: The Navrongo Community Health and FP Project is a quasi-experimental study designed to test the hypothesis that health and FP services can induce reproductive change in a traditional African societal setting. Just as in the case of Matlab, Navrongo had a history of less than favorable indicators of contraceptive knowledge as well as contraceptive use than corresponding indicators for the regions of southern Ghana. The Navrongo Community Health and Family Planning Project (CHFP) was launched in 1994 by the Navrongo Health Research Centre. Baseline characteristics of the population of women of reproductive age and their husbands were documented, and follow-up panel surveys and a longitudinal demographic surveillance system monitored changes in contraceptive knowledge, reproductive preferences, reproductive behavior, and fertility. Two experimental strategies were tested, each corresponding to two domains of the policy debate on the drivers of fertility decline. The first arm focused on the deployment of community volunteers. This arm tested the impact of mobilizing traditions to ensure sustainable volunteer participation in the program and community involvement in supervising and managing volunteer operations. Chiefs, lineage heads and women’s social networks were approached and then trained to build Community Health Compounds where nurses could be posted. Once this task was completed, this arm of the experiment focused on organizing the work of community health volunteers and building community participation for the management of their work.

The second arm of the experiment tested the impact of relocating nurses from sub-district clinics to community locations. Once nurses had been trained in community liaison methods and provided with motorbikes, basic drugs and primary health-care equipment, they were posted to Community Health Compounds where they offered vaccination services, treatment of common ailments (including malaria, acute respiratory infections and diarrhoeal diseases) and reproductive health and FP services. FP options
included the provision of injectable contraception, oral contraceptives and condoms in homes, and referral services for clinical methods, such as the five-year sub-dermal implant. Taken together, the two dimensions of the experiment comprised a four-cell design, since each dimension could be implemented independently, jointly or not at all. In the combined cell of the experiment, community liaison was directed to building community leadership of both volunteer and nurse service operations.

Early research showed that relocating nurses to communities increased the service volume, FP prevalence and immunization coverage, and expanded the range and quality of reproductive health care. Health services provided by a single nurse exceeded the typical case-load of a sub-district health centre. By 1999, district-wide experimental results showed that the total fertility rate had declined by one birth relative to comparison area levels; childhood mortality among 1–5-year-olds was reduced by over a third in the initial three project impact years that began in 1997 and by two thirds by the end of 2003. These results suggest that the Navrongo project achieved the child survival Millennium Development Goal in six years (Binka, Nazzar, and Phillips 1995; Pence et al. 2001; Phillips, Bawah, and Binka 2006; Pence et al. 2007). The program also had an impact on fertility. The arm of the experiment that focused exclusively on delivering health services to women and their children succeeded in reducing childhood mortality rates by half but had a negligible impact on fertility. The arm that focused on community-mobilization strategies and volunteer outreach however, led to a 15 percent reduction in fertility (Phillips, Bawah and Binka, 2005, 2006).

The evidence from Matlab as well as Navrongo can be interpreted as evidence that FP programs, when appropriately designed in keeping with the local conditions, can indeed be effective instruments of fertility decline and also produce many other effects on the well-being of women and their families.

5 Are FP Programs Cost-Effective?

The vast literature summarized in the sections above lead to the broad conclusion that FP programs have great direct and indirect benefits. The key question for policy-makers however, is how the costs of establishing and maintaining FP programs compare to other types of policy interventions. Advocates of FP programs have often argued that FP programs are one of the most cost-effective health interventions (Rhonda Smith et al. 2009). Others argue that the programs are prohibitively expensive and deliver little benefit since most declines in fertility are anyway induced by changes in desired family-size (Pritchett 1994a).

The debate on cost-effectiveness is at least partially driven by a lack of agreement on how to calculate costs and benefits. Estimation of costs can be challenging for several reasons. First, most available data pertains to average costs, rather than marginal costs. Since costs can vary depending on the level of scale attained by a program, comparison of average costs are problematic (Levine et al. 2006). Second, cost estimates vary based on the accounting method used, the method mix offered to couples, and the existing supply and service infrastructure (Dayaratna et al. 2000). Clinical costs are particularly sensitive to setting, given the broad differences in local salaries. Finally, most estimates of cost-effectiveness only calculate a program’s direct impact (couple year of contraception provided, births-averted, deaths-averted, etc.) while it is clear that such programs have long-term indirect effects as well, such as fewer
births in subsequent generations and/or improvements in overall health and well-being, and increased opportunities to reallocate time to tasks other than child-care.

Keeping in mind these challenges, Levine et al. (2006) synthesize information on costs of FP programs in a variety of settings. They estimate that the average cost per year for contraceptive supplies is about US$1.55, based on the existing mix of contraceptive methods used in developing countries. Program costs are higher, however, because they include health personnel and the cost of running facilities and outreach programs. These vary across regions and range on average from $2 to $35 per year of protection per person, depending on the mode of service delivery, such as social marketing, clinics, or community-based distribution. Costs in Africa tend to be higher, regardless of the service delivery mode, mainly because the existing supply of infrastructure is weaker than most other regions. The authors point out however, that this range of actual costs may in fact be narrower, because many existing programs are underutilized and economies of scale remain unutilized.

Similar estimates are also seen in Halperin, Stover, and Reynolds (2009) who examine the cost-structure of a variety of different programs. They find that the average cost per FP user is about US$20. The annual full service delivery cost per user however, ranges from $6 to $24 for temporary methods (pills, injectables and condoms), from $9 to $60 per acceptor of long-term methods (intrauterine device and implants), and from $30 to $100 per female sterilization.

The cost per year of contraceptives supplies does not provide adequate information on cost effectiveness of such an intervention. Most studies assume that in the absence of a program, users would simply not use contraception. The reality however, is that users may rely on the private sector or natural methods of FP to meet their need. Indeed, some evidence confirms this hypothesis. (Janowitz et al. 1992) for example, found that a Honduran social-marketing program distributed contraceptives to more than 40,000 couples but contraceptive prevalence remained unchanged. Users simply substituted one contraceptive brand or method for another. Similar results were found in Zaire by (Bertrand et al. 1993) who found that a community-based distribution program in then Zaire increased modern contraceptive prevalence without affecting total contraceptive prevalence. Users simply substituted modern methods for traditional ones.

Some researchers have examined the issue of cost-effectiveness by constructing cost-benefit ratios, where the benefits are measures such as number of pregnancies averted, number of births averted, number of deaths averted or disability-adjusted life-years (DALYs) that are attributed to the benefits of contraceptive use. The World-Bank’s Disease Control Priorities Project for example, estimates the cost of FP at $117 per DALY. Levine et al. (2006) summarize estimates from some major studies. These are summarized in Table 4. Note that the estimates vary significantly across regions, though there are large variations within regions that are masked by this analysis.

---

22 Information on the World-Bank’s Disease Control Priorities Project can be found at [http://www.dcp2.org/main](http://www.dcp2.org/main). The estimate of FP costs are constructed from 3 studies (not described in the database) that provide IUDs, voluntary sterilization, condoms and other barrier methods, implants, and oral contraceptives to a target population that includes all adult women of child-bearing age. The estimate of benefits presumably includes the impact of FP in lowering maternal and child mortality as well as morbidity.
It is important to note that the experimental studies discussed earlier in this paper, such as the Navrongo and Matlab projects, are typically more expensive than publicly funded programs. This is partly attributable to their smaller scale, the role of intensive data collection and intense focus on community outreach. Studies on Matlab, conducted between 1978 to 1985 argue that breakdown of costs is as follows: 60 percent for personnel and transportation, 8 percent for contraceptives, and 12 percent for various vaccines, oral rehydration salts, and general medicines, and service related supplies (G. B Simmons, Balk, and Faiz 1991). Costs per prevented birth are about $180 with a range of $150-$220. In the presence of indirect benefits or geographical spillovers however, this estimate would be expected to fall. Indeed, Simmons et al. even illustrated that the program’s cost was lower than the programs by the Bangladesh government, which had a larger administrative and bureaucratic apparatus.

Several recent studies also argue that the cost of FP programs can be lowered even further by linking such programs to established HIV programs. Halperin, Stover, and Reynolds (2009) for example, estimated that by offering FP services at voluntary counseling and testing sites in the 14 countries with high HIV prevalence, child HIV infections could be averted at a cost of US$489, and child deaths could be averted at a cost of US$278 per event—well below the costs of averting these events using traditional PMTCT services. In addition, these family programs would avert orphans at a cost of US$278, and maternal deaths at a cost of US$1,824 per event. In terms of DALY’s the benefits of such programs are even more startling: while HIV services cost approximately $27 per DALY, the reduction of mother-child transmission costs about $5 per DALY. Despite this evidence of cost-effectiveness however, FP is not currently included in most policies addressing mother-to-child transmission of HIV.

A final point on the issue of cost-effectiveness of FP programs is that most, if not all major studies, simply ignore the “spillover” benefits of such programs. To the extent that FP programs have now been shown to improve long-term health of women and their families, and have also been shown to be associated with greater female labor market participation of women and thus be poverty alleviating, the above-mentioned estimates should be regarded as lower bounds.23

In summary, we may conclude from this literature that even though it is often difficult to measure the correct costs of an FP program, and even harder to measure the effectiveness of such programs, most of the literature agrees that the programs are quite cost-effective and produce a plethora of benefits to the women who are targeted.

6 Family-Planning Programs Today

So far this paper has made two arguments: First, international interest in FP program has declined in the past twenty years; Second, the evidence on the benefits of a broad range of FP programs has accumulated over this same time period. This section argues that policy-makers may be well-advised to

23 Admittedly, modifying the cost-effectiveness framework to incorporate such benefits is quite difficult. A study in Thailand for example, used a multivariate regression framework to argue that the Thai FP program explained about 68 percent of the Thai fertility decline between 1972 and 1980 (Chao and Allen 1984) and this averted significant government expenditure. Even with a high discount rate, the estimated average monetary return on a dollar invested in the program is $16.31. The large number of assumptions implicit in this analysis necessitates caution, but the magnitude of the effect and the robustness of the result suggest that the benefits to a government from investing in effective FP programs may in many cases be positive.
consider establishing or scaling up FP programs, particularly in countries where fertility remains high and the demographic transition has not yet occurred.

**Is there a new window of opportunity?**

After nearly two decades of declining support, policy-makers appear to be once again turning their attention to FP programs. The landscape has changed significantly in the past twenty years. Many of the programs that were established early, i.e. during the 1960s and 1970s, appear to have taken root in their respective countries and remain domestically funded and managed. Some later programs, such as those in Bangladesh and Iran, also appear to have followed this course. In many other countries however, particularly those in Sub-Saharan Africa, programs have floundered in the era of declining international support.

A precise estimate in the decline in international support is difficult to measure because estimates of funding are generally grouped with all funding for reproductive health, which includes funding for HIV. A closer look at estimates from the UNFPA however, presents some interesting trends (Table 3). Note that in recent years, the amount of funding for reproductive health has indeed increased, but the percentage of funding allocated for FP has declined from 43 percent in 1998 to 6 percent in 2008 (UNFPA, 2009). Support has since then increased to nearly 10 percent in 2010 (UNFPA, 2011), suggesting that there is a renewed interest in such programs.

The renewal of interest in FP programs may also be driven by their ability to curb the spread of HIV. For the past decade, a great deal of funding and effort has been directed at combating HIV and other sexually transmitted diseases in many parts of the world, particularly in Sub-Saharan Africa. A significant number of new partnerships have been formed between international organizations, national governments, NGOs and civil society (Seltzer, 2002; Murray, Frenk and Evans, 2007; UNAIDS, 2008). Lessons from these experiences can combat some of the problems that plagued early FP programs in Sub-Saharan Africa: lack of donor coordination, management failures and service disruptions as a result of fragile health delivery systems. As discussed earlier, there is evidence that prevention of HIV through improved access to FP may be highly cost-effective (Bongaarts et al. 2008; Stover et al. 2006; Halperin, Stover, and Reynolds 2009). It may thus make sense to revisit the central themes of the Cairo agenda: integration and the provision of broad services.

**What should FP programs look like today?**

A policy-maker who is interested in FP programs will find that the programs implemented across the world in the past five decades have been diverse in their structure, implementation strategies, as well as their socio-economic contexts. These programs are now the basis of a large literature, only some of which has been reviewed in this paper. While this literature does not provide a “perfect recipe” or even

---

24 One example of this is the 2004 agreement between international donors, partners and national governments of the "Three Ones" principles -- One national AIDS Action Framework; One National AIDS Coordinating Authority; One agreed country-level Monitoring and Evaluation System—to achieve the most effective and efficient use of resources, and to ensure rapid action and results-based management (www.unaids.org). In a 2009 update of this strategy, it was noted that there was a marked improvement in the quality and greater technical support for the development of national AIDS strategies, the mainstreaming of AIDS into development plans, development of joint UN teams on AIDS and joint programmes, and development of accountability tools.
a list of “best-practices”, it does suggest that effective programs have a broad range of possible parameters.

A review of the literature does however provide a policy-maker with some very broad guiding principles. The first is that FP programs should never be regarded as a “substitute” for any other type of program focused on economic growth or investment in human capital. As we saw in section 3, FP programs seem to be more effective in societies that are making complementary investments in increasing female schooling, expanding labor-market opportunities, and experiencing economic changes that fundamentally change the cost-benefit tradeoff of high fertility. Even the contraceptives provided in experimental programs such as Matlab and Navrongo were most effective when coupled with other types of maternal and child health inputs. There is no doubt that the processes of socio-economic change can lower the demand for children in the long-run (Pritchett, 1994). But there is also little debate that the presence of an FP program can make it much easier for couples to attain their desired lowered fertility while those transitions are under way.

A second lesson for policy-makers is that investments in an FP program should be viewed much the same way as investments in any type of human-capital, particularly for women. Much research across the social sciences demonstrates the importance of investing in the human capital of women (Serageldin et al. 2001). The returns to other types of human-capital investments — such as education, labor-force participation and political participation --- have all been shown to have important private as well as social-returns almost everywhere in the world. Viewed through this lens, FP programs simply provide women with the opportunity to control fertility through access to modern, safe and effective methods of contraception. Given that it expands women’s choices and enables them to allocate their time to alternate activities, it is an investment in their ability to lead the life they have reason to value.

Another lesson for policy-makers is that even though the effects of an FP program may be small, and take time to attain, at the macro-level, these declines can nevertheless be significant. The 15% decline in fertility in the 70 villages of Matlab for example, together with a maternal and child health program, produced considerable spillover effects in the form of improved health for women and an improvement in the economic status of their families. Small declines in fertility at the micro-level can easily scale up to considerable numbers of individuals at the macro-level of a large population.

A final set of lessons for policy-makers concerns the structure of the programs themselves. These are summarized below:

1. **Programs should be voluntary:** Programs must be built around the goal of empowering individuals and families to make informed and responsible choices for themselves and their families. Evidence from even impoverished and conservative societies such as Bangladesh and Iran confirm that voluntary programs can be quite effective in the long-run.

2. **Programs should provide a full-range of services to women and adolescents:** Effective FP programs must recognize that individual’s needs differ based on their age and stage of life (which can include their marital status, housing status, and/or educational enrolment status). Adolescents may need reproductive health information and may wish to delay their first birth, younger married women
may wish to space their births, and older married women may wish to limit their births. Policies must differentiate between such individuals, both in terms of the services offered and the delivery systems themselves.

3. **Programs should mobilize and actively involve local communities:** Some of the best-performing FP and reproductive health programs have been owned and managed at the community level. Both the Matlab and Navrongo experiments highlight the importance of community mobilization strategies. Since community characteristics vary, as do attitudes towards fertility and reproductive health more generally, it is important that FP programs be carefully piloted and researched prior to roll-out or scale-up.

4. **Programs must be provided long-term commitment and coordinated funding:** Many examples in this paper illustrate that FP programs sometimes take several years to have demonstrable impact on levels of fertility. At the country-level coordinated long-term efforts have also been seen in not only East Asia, but also Bangladesh, Indonesia, Iran and Colombia (Seltzer, 2002; Robinson and Ross, 2007). At the project-level, the best examples of success – Matlab and Navrongo – have been interventions that have received long-term support from governments as well as funding agencies.

5. **An emphasis on evidence-based approach towards program design and implementation:** Since local conditions and the demand for children can vary significantly, any new FP program must be carefully piloted, scientifically launched and constantly evaluated. Evaluation must not only include an analysis of impacts on fertility, but on the ability to meet the needs of a community that it is launched in (Seltzer, 2002). We have already seen examples of success of such a policy in Taiwan and Bangladesh. It is important to note however, that in the past 50 years, there are many examples, from all across Africa, of micro-interventions that have increased the use of contraceptives in settings with formidable socio-economic and cultural barriers. Noteworthy among these are the Ghana Registered Midwives Project, the Ruhengeri Project in Rwanda, the Sudan Community-Based Family Health Project, and many others have been documented (McNamara, McGinn, Lauro and Ross, 1992; Segal, 1993). Many of these interventions could be re-examined for guidance on the next set of interventions.

**Conclusion**

This paper provides a brief overview of FP programs in the post WWII era, and a review of the academic literature on the effectiveness of these programs. They were heavily promoted in the 1960s and 1970s, but interest subsequently declined in international policy-circles. Academic evidence on the effectiveness of such programs has however, steadily accumulated over time. Researchers in this area face a series of methodological challenges: programs tend to be difficult to measure, they are placed non-randomly, and most of all, programs are implemented in societies that are experiencing rapid economic, social and political change, making it difficult to estimate causal relationships. Despite these challenges however, advances in data-collection and data-analysis have made it possible to study their impact. Estimates from cross-sectional data now suggest that FP programs may affect fertility in a variety of ways. They may for instance, increase the prevalence of modern contraception, or change social norms through information campaigns. There is no agreement on the magnitude of the impact of FP programs, but most researchers find that the effects are highest in countries that are experiencing economic growth and investments in human capital (particularly for women) more broadly. Panel
studies and experimental studies are able to better control for socio-economic trends and provide actual estimates of FP programs on fertility. Studies from Taiwan, Indonesia, Bangladesh and Ghana suggest that the magnitude of the effect may be between 5 and 20 percent. There is also evidence from several different countries such as Colombia and Bangladesh, that FP programs may have spillover effects in the long run, such as improvements in women's health, investments in human capital and employment in the labor force.

After nearly two decades, FP programs are once again being considered for international funding. Policy makers today have the opportunity to benefit from a large and new body of evidence on the design and implementation of population programs. While the literature does not provide a clear blue-print or recipe for an ideal FP program, some broad principles nevertheless emerge. This paper argues that FP programs should not be regarded as substitutes for any other type of policy aimed at increasing growth or development in a society. Rather, they should be regarded as a component of broader strategies to invest in the human capital (particularly of women) in a society. Programs should be voluntary, and should provide individuals with a broad range of supplies as well as support services that meet their needs. Programs should be established with the support and active involvement of the communities they are targeting. They should also be funded over long time-horizons. Programs should also be designed, implemented and evaluated with an evidence-based approach. Finally, instead of reducing fertility or promoting contraception, programs must be built around the goal of empowering individuals and families to make informed and responsible choices for themselves and their families. This freedom has intrinsic value for all human beings.
Table 1: The effects of socio-economic development (as measured by the “Development Index” variable) and the strength of FP programs (as measured by the “Program Effort” variable) on fertility decline between 1960-65 and 1980-85.

<table>
<thead>
<tr>
<th>Development Index</th>
<th>Strong</th>
<th>Moderate</th>
<th>Program effort</th>
<th>Weak</th>
<th>Very weak or none</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colombia</td>
<td>2.8</td>
<td>Fiji</td>
<td>Brazil</td>
<td>2.3</td>
<td>Iraq</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>3.5</td>
<td>Jamaica</td>
<td>Chile</td>
<td>2.5</td>
<td>Jordan</td>
</tr>
<tr>
<td>Korea Rep.</td>
<td>3.0</td>
<td>Malaysia</td>
<td>Costa Rica</td>
<td>3.4</td>
<td>Kuwait</td>
</tr>
<tr>
<td>Mauritius</td>
<td>3.3</td>
<td>Panama</td>
<td>Cyprus</td>
<td>1.1</td>
<td>Saudi Arabia</td>
</tr>
<tr>
<td>Mexico</td>
<td>2.6</td>
<td>Trinidad</td>
<td>Guyana</td>
<td>2.8</td>
<td></td>
</tr>
<tr>
<td>Singapore</td>
<td>3.2</td>
<td>Tobago</td>
<td>Venezuela</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>3.1</td>
<td>Average</td>
<td>Average</td>
<td>2.4</td>
<td>Average</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.8</td>
</tr>
<tr>
<td>Upper middle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>3.6</td>
<td>Dominican</td>
<td>Algeria</td>
<td>0.7</td>
<td>Bolivia</td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>1.9</td>
<td>Republic</td>
<td>Botswana</td>
<td>0.4</td>
<td>Congo</td>
</tr>
<tr>
<td>El Salvador</td>
<td>1.6</td>
<td></td>
<td>Ecuador</td>
<td>1.9</td>
<td>Nicaragua</td>
</tr>
<tr>
<td>Philippines</td>
<td>1.9</td>
<td></td>
<td>Peru</td>
<td>1.9</td>
<td>Paraguay</td>
</tr>
<tr>
<td>Thailand</td>
<td>2.9</td>
<td></td>
<td>Turkey</td>
<td>2.2</td>
<td>Syria</td>
</tr>
<tr>
<td>Tunisia</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>2.7</td>
<td>Average</td>
<td>Average</td>
<td>2.4</td>
<td>Average</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.8</td>
</tr>
<tr>
<td>Lower middle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indonesia</td>
<td>1.3</td>
<td></td>
<td>Egypt</td>
<td>1.8</td>
<td>Burma</td>
</tr>
<tr>
<td>Guatemala</td>
<td>0.7</td>
<td></td>
<td>Ghana</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>Honduras</td>
<td>1.2</td>
<td></td>
<td>Ivory Coast</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>Kenya</td>
<td>0.0</td>
<td></td>
<td>Lesotho</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Liberia</td>
<td>−0.2</td>
<td></td>
<td>Nigeria</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>Morocco</td>
<td>1.7</td>
<td></td>
<td>Togo</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>Papua N.G</td>
<td>0.6</td>
<td></td>
<td>Zaire</td>
<td>−0.1</td>
<td></td>
</tr>
<tr>
<td>Tanzania</td>
<td>−0.2</td>
<td></td>
<td>Zambia</td>
<td>−0.6</td>
<td></td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>1.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>1.3</td>
<td>Average</td>
<td>Average</td>
<td>0.8</td>
<td>Average</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>1.1</td>
<td>Bangladesh</td>
<td>Gambia</td>
<td>0.0</td>
<td>Benin</td>
</tr>
<tr>
<td>Haiti</td>
<td>1.1</td>
<td></td>
<td>Burkina Faso</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>Pakistan</td>
<td>0.0</td>
<td></td>
<td>Burundi</td>
<td>−0.8</td>
<td></td>
</tr>
<tr>
<td>Rwanda</td>
<td>−0.8</td>
<td></td>
<td>Central African</td>
<td></td>
<td>Republic</td>
</tr>
<tr>
<td>Senegal</td>
<td>0.2</td>
<td></td>
<td>Ethiopia</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Guinea</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Madagascar</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Malawi</td>
<td>−0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mozambique</td>
<td>−0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Niger</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sierra Leone</td>
<td>−0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sudan</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Uganda</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Yemen, N.</td>
<td>−0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Yemen, PDR</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>1.1</td>
<td>Average</td>
<td>Average</td>
<td>0.5</td>
<td>Average</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0</td>
<td></td>
</tr>
</tbody>
</table>

Notes: (i) The development index is based on 4 variables: GNP per-capita, the infant mortality rate, gross enrollment rate, and a composite measure of communication (number of TVs, radios and cars per capita). Each variable had a score between 0 and 100, so the aggregate index has a range of 0-400. Countries were classified as follows: “High” (value: 225-400), “Upper-Middle” (value: 175-224); “Lower-Middle” (value: 125-174); “Low” (value: 0-124); (ii) Program effort consists of 30 separate measures grouped into four components: policy, service, record keeping, and availability of methods. Scores ranging from 0-4 were calculated for each of the 30 items. The maximum possible total score was 120. Countries were classified as follows: “Strong” (value: 67-120), “Moderate” (value: 46-66), “Weak” (value: 21-45), “Very Weak” (value: 0-20).

Table 2: Selected Characteristics of True Experimental Studies of Family Planning Program Effects

<table>
<thead>
<tr>
<th>Country (Year)</th>
<th>Author(s)</th>
<th>Program</th>
<th>Author Conclusion</th>
<th>$r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil (1981)</td>
<td>Forei &amp; Forei</td>
<td>Supervisory visit frequency</td>
<td>No effect for number of new clients, returns</td>
<td>0.09</td>
</tr>
<tr>
<td>Colombia (1969)</td>
<td>Simmons</td>
<td>Home visit, mailed pamphlet</td>
<td>No effect for clinic attendance</td>
<td>0.04</td>
</tr>
<tr>
<td>Colombia (1979–1981)</td>
<td>Gomez</td>
<td>Health promoter distribution</td>
<td>Increased oral contraceptive use</td>
<td>0.09</td>
</tr>
<tr>
<td>Hong Kong (1966)</td>
<td>Population Council</td>
<td>Home visit, visitor qualification</td>
<td>Increased clinic attendance</td>
<td>0.14</td>
</tr>
<tr>
<td>Hong Kong (1968/69)</td>
<td>Chan</td>
<td>Home visit</td>
<td>No effect for IUD retention</td>
<td>0.02</td>
</tr>
<tr>
<td>Korea (1966/67)</td>
<td>Yang</td>
<td>Mother education class</td>
<td>IUD prevalence increased in low prevalence areas; termination decreased with special education</td>
<td>0.02</td>
</tr>
<tr>
<td>Korea (1967)</td>
<td>Bang</td>
<td>Early clinic return schedule</td>
<td>Increased IUD removal and expulsion</td>
<td>-0.08</td>
</tr>
<tr>
<td>Mexico (1968/87)</td>
<td>Macorra et al.</td>
<td>Supermarket condom location</td>
<td>Increased condom sales</td>
<td>...</td>
</tr>
<tr>
<td>Nepal (1973)</td>
<td>Gubhaju et al.</td>
<td>Single-purpose field worker</td>
<td>Increased pill continuation</td>
<td>0.08</td>
</tr>
<tr>
<td>Peru (1985–1987)</td>
<td>Zambrano et al.</td>
<td>Physician frequency at clinic</td>
<td>Increased visits and IUD insertions</td>
<td>...</td>
</tr>
<tr>
<td>Philippines (1973)</td>
<td>Phillips et al.</td>
<td>Motivator incentive and at-large affiliation</td>
<td>Increased motivator performance</td>
<td>0.16</td>
</tr>
<tr>
<td>Sri Lanka (1983–1985)</td>
<td>Vidyasagara et al.</td>
<td>Satisfied user–midwife team</td>
<td>Increased IUD acceptors, no effect for termination</td>
<td>0.14</td>
</tr>
<tr>
<td>Taiwan (1964)</td>
<td>Freedman &amp; Takeshita</td>
<td>Home visit, wife vs couple involved, mailing</td>
<td>Home visit increased contraceptive acceptance; wife or couple involvement equally effective; mailing ineffective</td>
<td>0.05</td>
</tr>
<tr>
<td>Taiwan (1971)</td>
<td>Chang et al.</td>
<td>Field worker incentive</td>
<td>Increased contraceptive acceptance</td>
<td>0.07</td>
</tr>
<tr>
<td>Thailand (1969/70)</td>
<td>Rosenfield &amp; Limcharoen</td>
<td>Midwife or physician pill prescription</td>
<td>Increased acceptance and continuation with midwife prescription</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Note: IUD = intrauterine device.

*aCoefficient could not be calculated from information provided.

Source: Bauman (1997)
### Table 3: Final donor expenditures for population-assistance, by category of population activity (in $ and percentages)

**Source:** UNFPA (2008)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Family planning services</strong> (Millions of current $US)</td>
<td>722.83</td>
<td>612.4</td>
<td>517.9</td>
<td>605.4</td>
<td>725.3</td>
<td>404.7</td>
<td>445.2</td>
<td>500.9</td>
<td>393.5</td>
<td>461.8</td>
<td>572.4</td>
</tr>
<tr>
<td><strong>Basic reproductive health services</strong> (Millions of current $US)</td>
<td>369.82</td>
<td>496.5</td>
<td>516.4</td>
<td>492.3</td>
<td>781.9</td>
<td>1090.6</td>
<td>1036.9</td>
<td>1135.9</td>
<td>1478.3</td>
<td>1467.1</td>
<td>1716.8</td>
</tr>
<tr>
<td><strong>Sexually transmitted diseases and HIV/AIDS activities</strong> (Millions of current $US)</td>
<td>336.20</td>
<td>380.7</td>
<td>575.0</td>
<td>793.7</td>
<td>1339.2</td>
<td>1862.2</td>
<td>2754.8</td>
<td>4884.8</td>
<td>5102.2</td>
<td>6540.2</td>
<td>7702.0</td>
</tr>
<tr>
<td><strong>Basic research, data and population and development policy analysis</strong> (Millions of current $US)</td>
<td>252.15</td>
<td>182.1</td>
<td>162.7</td>
<td>156.4</td>
<td>315.6</td>
<td>489.4</td>
<td>576.3</td>
<td>278.3</td>
<td>342.4</td>
<td>297.7</td>
<td>399.1</td>
</tr>
<tr>
<td><strong>Total activities</strong> (Millions of current $US)</td>
<td>1,681</td>
<td>1,655</td>
<td>1,781</td>
<td>2,051</td>
<td>3,162</td>
<td>3,847</td>
<td>4,813</td>
<td>6,800</td>
<td>7,318</td>
<td>8,767</td>
<td>10,391</td>
</tr>
</tbody>
</table>

- Percentages have been rounded off and may not add to 100 per cent.
- The development banks are not included in the final expenditures shown, as the banks’ loan agreements are often disbursed over several years.
- 2000 data differ from the figures in the 2000 report, due to additional information received.
- Distribution for Germany has been partially estimated based on 2001 percentages. Distribution for Luxembourg has been estimated based on 2001 data. Distribution for Italy has been estimated based on 2000 data.
- Distribution for the European Union has been estimated by NIDI based on data from the European Commission and the DAC Watch of the European Union, IPPF, January 2002.
- 2002 data differ from the figures in the 2002 report, due to additional data received.
- Distribution for France has been estimated by NIDI based on data from the European Commission and the DAC Watch of the European Union, IPPF, January 2002.
- Basic reproductive health care services for Sweden included family planning services.
- Basic reproductive health care services for the United Kingdom included family planning services.
Table 4: Average costs per benefit of Family Planning

<table>
<thead>
<tr>
<th>Region</th>
<th>Births averted</th>
<th>Infant deaths averted</th>
<th>Maternal deaths averted</th>
<th>Disability-adjusted life years saved</th>
<th>Years of life lost averted</th>
<th>Years lived with disability</th>
</tr>
</thead>
<tbody>
<tr>
<td>East Asia and the Pacific</td>
<td>163</td>
<td>4,907</td>
<td>12,880</td>
<td>60</td>
<td>110</td>
<td>103</td>
</tr>
<tr>
<td>Latin America and the Caribbean</td>
<td>87</td>
<td>2,316</td>
<td>34,564</td>
<td>53</td>
<td>66</td>
<td>167</td>
</tr>
<tr>
<td>Middle East and North Africa</td>
<td>57</td>
<td>1,999</td>
<td>18,917</td>
<td>49</td>
<td>55</td>
<td>209</td>
</tr>
<tr>
<td>South Asia</td>
<td>113</td>
<td>1,577</td>
<td>5,772</td>
<td>30</td>
<td>37</td>
<td>98</td>
</tr>
<tr>
<td>Sub-Saharan Africa</td>
<td>131</td>
<td>1,807</td>
<td>10,231</td>
<td>34</td>
<td>37</td>
<td>184</td>
</tr>
</tbody>
</table>

Source: Afton, based on a model by APS and others (2000).

Note: The model used country-level data for 68 developing countries. Output costs were based on Population Action International estimates in 1984 of the public sector cost per user. Estimates are not available for Eastern Europe and Central Asia.

Source: Levine et al. (2006)
REFERENCES


countries. New York: Oxford University Press.
Sinha, N. 2005. Fertility, child work, and schooling consequences of family planning programs: Evidence from an experiment in rural


