The Impact of Population Policies: Reply

LANT H. PRITCHETT

I WOULD LIKE TO GIVE a glimpse of the forest before discussing the specific trees selected for the ax by James Knowles, John Akin, and David Guilkey (1994) and by John Bongaarts (1994a). The six points made in my article are three facts and three supports.

1 Empirically, nearly all the differences across countries in actual fertility are explained by cross-country differences in women's desired fertility. Desired fertility plus a constant is an excellent prediction of actual fertility.

2 The two most potent objections to the use of desired fertility as an explanatory variable can be overcome by using both retrospective (was a previous birth wanted?) and prospective (is another birth wanted?) data on desired fertility together with the statistical estimation technique of instrumental variables. Neither ex-post rationalization of prior births nor the influence of contraceptive costs on reported fertility desires determines the empirical results.

3 In explaining differences in fertility such measures as contraceptive use, "unmet need," and family planning effort have an empirically small (though at times statistically significant) impact and explain very little of fertility variation once differences in desired fertility are accounted for.

The three supports of these facts showed that they were consistent with economic common sense, consistent with much of the literature, and not refuted by arguments commonly adduced in support of the presumed importance of family planning programs.

4 A small effect of family planning programs is intuitively plausible when the cost of avoiding a child is compared with the cost of having a child. Having a child is just too important and expensive for the decision to be strongly affected by contraceptive costs.

5 A small family planning program effect on fertility is consistent with historical experience, which shows that modern contraceptive technology

is not a precondition for fertility decline, and with the household literature on the impact of contraceptive access on fertility.

6 Arguments based on the existence of "unmet need" for contraception, the speed of the demographic transition, or experimental family planning programs do not provide contradictory evidence to an empirically small family planning program impact.

Their self-assured tone notwithstanding, neither the objections raised by Knowles, Akin, and Guilkey (KAG) nor those raised by Bongaarts present a serious challenge to any of these six points.

Knowles, Akin, and Guilkey

The first point KAG make (one whose importance, they say, "cannot be overstated") is that excess fertility, the difference between actual and desired fertility, does not explain any variation in fertility. I agree with this point, not surprisingly, since I said it first ("Excess fertility is not systematically related to the level of fertility," Pritchett 1994: 5). This fact about excess fertility is a simple consequence of point 1 above, that desired fertility explains almost all of actual fertility variation, leaving very little to be explained by excess fertility.

However, while KAG agree with my facts they argue that because the desired fertility rate (DTFR) (or the wanted fertility rate [WTFR]) is a "slight mutation" of the total fertility rate (TFR), my article is "seriously flawed" by its use of "an essentially tautological model." Their claim is that if one variable (say, DTFR) is derived from another variable (say, TFR), the finding of a strong correlation is a statistical artifact. Although this may sound plausible,[1] as a general proposition, and in this particular instance, their claim is completely false. KAG's mistake stems from a confusion of a necessary (or definitional) fact (that DTFR is a mutation of TFR) with an empirical, contingent fact (that the mutation is slight).

The easiest way to see that KAG's basic claim is false is to consider their own empirical evidence. Excess fertility (ExTFR) is constructed from actual TFR in exactly the same way DTFR is: by deleting some births from TFR—namely, those births that happen to be wanted. DTFR and ExTFR are therefore exact equivalents: they are both "slight mutations" of TFR. But as KAG show, ExTFR has no statistical relationship with TFR at all, flatly contradicting their claim that any independent variable derived from some "slight" transformation of the dependent variable must have a strong relationship with the dependent variable. What distinguishes DTFR and ExTFR is not some necessary fact about their definition or construction but the empirical fact about the fraction of births women report as wanted or unwanted. Although KAG dismiss the high explanatory power of DTFR as "stemming from the inclusion of a slight transformation of the dependent

variable," the question of whether or not the transformation of deleting unwanted births was empirically "slight" is exactly what is under consideration. Simply because the correlation was in fact high does not mean that the correlation was necessarily high. If DTFR were constant across countries—say if all women everywhere wanted 2.1 children—then the empirical results would have been the opposite. ExTFR would have high explanatory power and DTFR would be unrelated to fertility. No particular result was built into the data.[2]

So while I agree that the mutation of TFR to DTFR is empirically slight (see point 1 above), there is nothing artifactual about that empirical fact. Once the confusion between necessary and empirical facts is clarified, the rest of KAG's comment, although technically competent, is completely beside the point.[3] To use KAG's analogy to precipitation,[4] I am perfectly happy to admit that the empirical data show that excess fertility is the hail of precipitation. But if this fact about fertility is so obvious as to not "advance the science of meteorology," why (metaphorically) do so many international conferences on precipitation spend so much time discussing hail and so little discussing rainfall?

Bongaarts

On the key point just discussed, John Bongaarts (1994a) completely disagrees with KAG. He does not agree that excess (unwanted) fertility is just a *small* random error around actual fertility. He is at pains to stress that the level of unwanted fertility is *large*. I have little disagreement. As an analogy, suppose that people in the United States were surveyed about their desired weight and everyone, regardless of his or her actual weight, would like to weigh 10 pounds less. We could then say things like "undesired weight in the USA is nearly one million tons."[5] Such a finding about the magnitude of unwanted weight, though interesting, would not explain why some people were heavy and others light. Nor would the fact that the weight was undesired imply that there existed any diet plan that would rid Americans of their unwanted weight. All of the fertility regressions I presented have a significant constant term—which gives the average level of unwanted fertility. However, this constant term is exactly that, a *constant*. Countries move from high fertility to low fertility not because *unwanted* fertility goes down, but because *desired* fertility goes down. Whatever the magnitude of unwanted fertility, family planning programs aimed at lowering fertility may have empirically the same effect on fertility as most diets have on weight: none.

In judging the impact of family planning programs on fertility, Bongaarts accuses me of my "most egregious error" in using incremental R^2 to judge the importance of an explanatory variable. For good measure, he re-

fers me to "standard statistical textbooks." But, from a standard statistics textbook one gets a standard textbook understanding of statistics. If one understood the reasons *why* incremental R^2 is not appropriate as a mechanically applied statistical procedure, one would also understand why its use is perfectly appropriate in my case. Bongaarts's "counter-example" is not a counter-example to my use of incremental R^2, just an example of poor usage.

The incremental R^2 is inappropriate for judging the relative importance of a set of independent variables when no causal model is specified among those variables. The attribution of the explanatory power for the dependent variable resulting from the covariation among the independent variables to any given independent variable based on the sequence of inclusion in the estimation is statistically arbitrary. However, this general injunction against using incremental R^2 as a statistical procedure applies only when the causal relationships among the independent variables are not known. But that causal structure is exactly what, as I argued extensively (Pritchett 1994: 10–12), we *do* know in this instance. Lower desired fertility (taken as the number of children women would have if perfect contraception were free) causes contraceptive prevalence (CPV), but contraceptive prevalence does *not* cause lower desired fertility. Hence, in his "counter-example" Bongaarts does indeed present an instance of an illegitimate use of incremental R^2. Judging the impact of female education on fertility by its incremental impact after controlling for CPV, which is caused by female education, *is* wrong. However, unless Bongaarts cares to argue that *current* contraceptive prevalence affects the decisions in the *past* that resulted in the current stock of female education (and I think some law of physics actually prevents this), the causal structures of his example and the use of incremental R^2 in my article are clearly different. Judging the magnitude of the impact of CPV (or family planning effort [FPE]) after controlling for a variable it does not cause (DTFR) is not poor use of incremental R^2. The reason the two uses are different is that although the two cases are *statistically* equivalent they are not *analytically* equivalent because of our prior, nonstatistical knowledge about the structure of causation.[6]

Bongaarts does some verbal theorizing about emerging desires to limit fertility and the impact of family planning programs. He suggests that this theorizing leads to a new, different regression ("differences [between our results] are that unwanted fertility is the dependent variable . . . "). Bongaarts suggests that this "new" regression supports his theory. He appears not to realize that his "new" regression is *exactly* my regression, just in a different algebraic form that has no impact at all on the estimated coefficient. Subtracting WTFR from both sides of the regression (so that unwanted fertility is the dependent variable) does absolutely nothing to the coefficient on FPE.[7] The *only* difference between Bongaarts's reported regressions and those in my article is the difference in the sample.

In fact, Bongaarts has found one particular sample and one particular desired fertility measure for which he argues FPE has a large coefficient.[8] How robust is the one regression he reports to reasonable variations in sample or variable definitions? Table 1 presents some variants with different samples (1970s and 1980s separately) and variables (DTFR as well as WTFR). The result given in the top row of the table is the one regression Bongaarts reports. Immediately below it are the results for other samples and variables that I calculated with the data I had available. With my sample the results that best support Bongaarts's case do appear for the 1980s. In the 1980s the coefficient is about –.020 (the estimates are –.019 [DTFR] and –.021 [WTFR]), and using the incremental R^2 FPE explains between 3.5 and 4.5 percent of actual fertility variations. On the other hand, FPE had little or no impact in the 1970s. I think a safer conclusion, one that better represents the data, is that the estimated FPE effect varies from as low as .005 to as high as .028 in one particular sample. Settling on .02 as an estimate of the FPE effect (rather than picking from all estimates the one that happens to be the largest) seems fair—and is actually overly generous given the potentially serious problems with the endogeneity of FPE.[9]

TABLE 1 Regressions of family planning program effort (FPE) and desired or wanted fertility (DTFR, WTFR) on total fertility (TFR)

Coefficient on FPE	Coefficient on DTFR or WTFR	Incremental R^2	N	Fertility measure	Description of sample
Bongaarts's reported result (1994a)					
–.028 (5.1)	.677 (9.44)		25	WTFR	Data from 1980s, using 1982 FPE data
Other possible estimates of the FPE effect					
–.0058 (2.21)	.828 (20.0)	.006	65	DTFR	All data
–.019 (4.19)	.670 (10.53)	.036	31	DTFR	Data from 1980s, using 1982 FPE data
–.0048 (1.55)	.819 (12.87)	.005	34	DTFR	Data from 1970s, using 1972 FPE data
–.0105 (3.31)	.753 (11.72)	.033	46	WTFR	Data from entire sample, using FPE data most nearly preceding fertility data
–.0208 (3.27)	.676 (8.32)	.045	15	WTFR	Data from 1980s, using 1982 FPE data
–.0088 (2.29)	.753 (7.68)	.033	31	WTFR	Data from 1970s, using FPE data from 1972

NOTES: *t*-statistics in parentheses. The reported results for WTFR in Bongaarts 1994a simply add 1 to his reported coefficient and then divide by the standard error implied by his reported *t*-statistic (the standard error is unchanged by the transformation).

SOURCE: For data and data sources, see Pritchett 1994.

Even if we accept the best of all possible results for the importance of FPE, what does this disprove? I feel perfectly comfortable saying that a variable with explanatory power of less than 5 percent of the fertility variance and with a coefficient of .02 (or even .028) can be safely described as having a demonstrable, but quantitatively small, impact. An incremental effect of .02 would require a 50-point (on a scale of 0 to 100) increase in FPE to reduce fertility by just one birth. While Bongaarts takes the fact that a 50-point increase in FPE could reduce fertility by (in his estimate) 1.4 births as evidence of the efficacy of family planning programs, he does not point out that a 50-point increase is enormous and improbable—indeed, in most cases impossible. First, a move of 50 points means a change from a practically nonexistent program like that of Chad (20), Mauritania (21), or Bhutan (22) to one like that of Bangladesh (72). Second, only one country (Botswana) actually did improve FPE by 50 points over the seven years from 1982 to 1989. Third, the practical maximum for FPE is around 80 percent of the theoretical maximum effort (that is where Korea, Taiwan, Thailand, Sri Lanka, and Indonesia are clustered). Even China achieves only 87 percent of the maximum—and I doubt Bongaarts intends to advocate a Chinese-style program. Taking 80 as the practical voluntary maximum program effort score, most countries do not score below 30 (only 28 of 98 countries scored are below 30); thus achieving a reduction of even one birth by a 50-point increase in FPE is rarely possible.

What then is the scope for fertility reduction through increased family planning effort, even assuming the generous estimates of the effect based only on data from the 1980s? What if in 1989 (the year with the latest available data) every country in the world had a program as good as the very best countries and achieved an FPE rating of 80? Granted this is a ludicrous assumption given the practical, administrative, and financial constraints—but what if it did happen? If every country in the world were to have the strongest observed voluntary family planning program, the developing country weighted average of fertility in 1989 would fall by just 8 percent.[10] Most of the populous countries (China, India, Bangladesh, Indonesia) already have strong programs. Even excluding China (which looms large in world calculations and which has a zero assumed effect since FPE is already higher than the practical maximum for a voluntary program), the world fertility fall is only 10 percent—from 4.5 to 4.1 births. (Even at Bongaarts's generous estimate of .028, the fertility fall is just 11 percent with China and 14 percent without.) This maximum possible fertility reduction achievable through family planning effort of 0.4 births (at an FPE coefficient of .02) is usefully contrasted with the observed fall in the weighted average fertility in low- and middle-income countries since 1965 of 2.8 births—seven times greater than the largest conceivable incremental impact of family planning programs (World Bank 1994).

The larger issue

The question of whether these family planning program effects are quantitatively small or not returns the discussion to the larger issue at hand. The issue is the relative importance for the demographic transition of the "supply" of contraceptives, which includes all aspects of family planning programs that affect individuals' cost of contracepting, and the demand for children. I argued that, although of course both may play some role, the reality of fertility declines is that it is much more like a 5–95 supply–demand split than the 50–50 (or in some cases 95–5) division that much of the older (and even much of the new, more balanced) family planning literature seems to convey.

Referring to this last point, Bongaarts asserts that "much of the article attacks a straw man." To this I have three responses. First, if the view that my article argued against, that contraceptive supply and family planning programs play a large independent role in the demographic transition, is a straw man then this must be the land of Oz, for this particular straw man seems to be alive and dancing. For every point raised I provided direct quotes from family planning advocates, generally from widely cited publications. The approach of the Cairo conference is bringing many more examples to hand. For instance, Sinding, Ross, and Rosenfield (1994) assert that "improved access to high-quality reproductive health services . . . will carry the world a very long way toward replacement-level fertility" (p. 27). Simply because family planning program advocates have now learned to twin references to contraceptives with references to social conditions does not mean the true relative empirical magnitudes of the two effects have been fully appreciated.

Bongaarts's own article in *Science* (1994b) starts a list of recommended population policies with "1) Reduce unwanted pregnancies by strengthening family planning programs" (p. 773). Moreover, Figure 4 in the same article suggests that 1.9 billion in population increase by the year 2100 will come from "Unwanted fertility," nearly twice the 1 billion increase attributed to "High desired family size." While a careful reading of the *Science* article can detect revisionism relative to more outlandish claims for the importance of family planning programs, I think the average educated reader could easily confuse Bongaarts's nuanced view of population policies with the "straw man" view that access to low-priced contraceptives to eliminate unwanted fertility is the most important (if not the only) key to slowing population growth. The casual or naive reader might fairly conclude from Bongaarts's Figure 4 that eliminating unwanted fertility will be twice as important in affecting future population growth as reducing desired family size. Given the historical record and recent experience, such an estimate would be wrong by roughly an order of magnitude.

Second, even if it were true that I was attacking a straw man, the public burning of a straw man can be of cathartic value, especially when its frame closely resembles skeletons in one's closet. The straw man view enjoys wide currency among government officials, nonspecialist intellectuals, journalists, and the lay public. One suspects these straw man beliefs did not spring into their minds fully formed, but rather took some considerable encouragement from certain demographers and family planning program advocates. I quoted Paul Kennedy in my original article stating exactly the straw man view not to imply that he was egregiously uninformed but rather the opposite: because he had been informed by the literature. Professor Kennedy, a well-known historian at Yale, is much smarter than the average nonspecialist and yet after reviewing the population and family planning literature he came to exactly the straw man conclusion. This is a telling indication of what that literature is saying to the world. I certainly welcome Bongaarts's support in revising sharply downward the existing, overly optimistic beliefs about the importance of contraceptive access for the course of the demographic transition and in emphasizing the importance for fertility decisions of the economic and social conditions faced by women in developing countries.

Third, although I may have unnecessarily devoted space to clarifying arguments already well understood by some, I also attacked the strongest evidence I could find that might be considered as contradicting the view I articulated. An article that addresses the fertility impact of access-induced increases in contraceptive prevalence, family planning effort, and "unmet need" *and* discusses the implications of the Matlab experience *and* explains episodes of rapid fertility decline can be accused of being too long, but cannot be accused of failing to address the critical issues. Those who were already convinced that the "family planning view" was wrong might have indeed felt I was attacking a straw man. However, I would hope that others have become so persuaded as a result of reading my article.

Notes

1 Actually I do not find it plausible, but I am inferring that it must sound plausible from the fact that others have raised essentially this same point in personal communication with me (Bhushan and Kincaid 1994).

2 That nothing is built into the data is also revealed by the fact that using women's average ideal number of children (AINC)—which is not derived from the same data on births as TFR—gives roughly the same results (once corrected for classic errors-in-variables bias by using the statistical technique of instrumental variables estimation).

3 KAG show that in the model $TFR = \beta \times DTFR + \varepsilon$, if β is, by assumption, exactly 1, then random measurement error (perhaps compounded by truncation) would bias the estimate of β downward. With a small amount of measurement error I could have obtained my reported empirical results even if $\beta = 1$ exactly. While correct, I cannot understand why they would raise this issue.

If one accepts that $\beta = 1$ and that the rest of the differences between actual and desired fertility are just small random error, this downward bias reinforces the thrust of my article rather than challenges any part of my analysis.

4 KAG's own reference to restricting the precipitation–rainfall relationship to warmer climates acknowledges that they are dealing with an empirical, not tautological, fact when they assert that the precipitation–rainfall relationship is trivially true. While the precipitation–rainfall relationship is perhaps trivial in North Carolina (in Raleigh snow is less than 2 percent of precipitation), as a native Rocky Mountain Westerner I am more conscious of the fact that snow is 39 percent of precipitation in Denver and Great Falls, 38 percent in Cheyenne, 36 percent in Salt Lake City, and 31 percent in Spokane (US Bureau of the Census 1993).

5 Ten pounds for each of the 197,046,000 individuals over age 15 in the US comes to 985,230 tons.

6 In terms of equations the model is:

$$TFR = \beta_1 \times CPV + \beta_2 \times FE,$$
$$CPV = \alpha \times FE,\ \alpha > 0,$$
$$\text{but } FE = \delta \times CPV,\ \delta = 0$$

(where FE is female education). In this case, since CPV does not cause FE, adding CPV to the fertility equation after FE is appropriate, but not vice versa since FE does cause CPV. Again, the point is that the full structural model matters for the judgment about whether the incremental R^2 is an appropriate statistical technique.

7 That is, if one has the model: $TFR = \beta \times WTFR + \delta \times FPE + \varepsilon$ and then subtracts WTFR from both sides, the new model is: $(TFR - WTFR) = (\beta - 1) \times WTFR + \delta \times FPE + \eta$. The coefficient on FPE is unchanged (numerically equivalent) and the coefficient on WTFR is simply the original coefficient less one.

8 It might also be pointed out that Bongaarts takes the liberty in Figure 1 of his Comment of reporting only the part of his sample that best supports his case. Although the regression he presents has a sample of 25, the figure shows only 13 countries. For those countries alone, the relationship appears quite close (the R^2 for unwanted fertility in the figure shown is over 0.7 just for FPE alone). In contrast, in the regression with the full sample the R^2 was 0.56 for FPE *and* WTFR together. Clearly the UnWTFR–FPE fit is much less tight and striking for the countries *not* shown than for those selected for display.

9 The results do prompt a question about whether the FPE effect is overestimated, a question I only touched on briefly in my article because the estimated FPE effects were so small in any case (see p. 19). To the extent that FPE is determined by the intensity among the population of the desire to limit fertility, which it certainly is to at least some degree, the FPE coefficient (even controlling for DTFR) will be overstated (a result consistent with Schultz 1994).

10 This and later simulations are done by calculating country by country (for the 88 countries for which both FPE and TFR data were available) what fertility would be if FPE were 80 (using the formula $TFR_{FPE=80} = Actual\ TFR - \beta \times (80 - Actual\ FPE)$ where β (<0) is the coefficient on FPE in the fertility regression. The population-weighted average fertility (using 1989 populations) for acutal TFR and simulated TFR at FPE = 80 are compared. Since China is above 80 the effect for China is set to zero (although it is plausible that fertility would rise if China were to have an entirely voluntary program). Also, it is assumed that increases in FPE do not drive fertility lower than replacement level (this does not greatly affect the calculation in levels, as it would happen only in a few small countries).

References

Bhushan, Indu and D. Lawrence Kincaid. 1994. "Desired fertility and the impact of population policies: A rejoinder," personal communication.

Bongaarts, John. 1994a. "The impact of population policies: Comment," *Population and Development Review* 20, no. 3: 616–620.

———. 1994b. "Population policy options in the developing world," *Science* 263 (11 February): 771–776.

Knowles, James C., John S. Akin, and David K. Guilkey. 1994. "The impact of population policies: Comment," *Population and Development Review* 20, no. 3: 611–615.

Pritchett, Lant H. 1994. "Desired fertility and the impact of population policies," *Population and Development Review* 20, no. 1: 1–55.

Schultz, T. Paul. 1994. "Sources of fertility decline in modern economic growth: Is aggregate evidence on the demographic transition credible?" mimeographed, Yale University (January).

Sinding, Steven W., John A. Ross, and Allan G. Rosenfield. 1994. "Seeking common ground: Unmet need and demographic goals," *International Family Planning Perspectives* 20, no. 1: 23–27, 32.

US Bureau of the Census. 1993. *Statistical Abstract of the United States: 1993.* (113th edition). Washington, D.C.: US Government Printing Office.

World Bank. 1994. *World Development Report 1994.* New York: Oxford University Press.