14075

THE WORLD BANK ECONOMIC REVIEW, VOL. 7, NO. 3: 319-348

Sept. 1993

# The Determinants and Consequences of the Placement of Government Programs in Indonesia

Mark M. Pitt, Mark R. Rosenzweig, and Donna M. Gibbons

*Most estimates of the consequences of public programs rely on the cross-sectional association between region-specific programs and program outcomes. Such estimates assume that the spatial distribution of programs is random. This article reports estimates of the effects of public programs on basic human capital indicators and the biases in conventional cross-sectional estimates of program effects due to non-random program placement. The estimates are obtained from pooled observations on human capital outcomes, socioeconomic variables, and program coverage at the* kecamatan *(subdistrict) level. The observations are based on successive sets of Indonesian cross-sectional household and administrative data during 1976–86. The determinants of the spatial allocation of programs in Indonesia in 1976–86 are also estimated.*

*The empirical results indicate that the presence of grade and middle schools in villages has a significant positive effect on the school attendance rates of teenagers. The presence of health clinics in villages also positively affects the schooling of females ages 10–18. However, no evidence is found of any significant effects of the presence of family planning and health programs on either the survival rates of children or on cumulative fertility. The estimates also suggest that the use of cross-sectional data results in substantial biases in the estimates of program effects because of the evident nonrandom spatial allocation of public programs.*

Developing countries invest heavily in a wide variety of social sector programs, with health, fertility control, and schooling being central among them. Much literature in the social sciences is devoted to evaluating these programs. Most such studies have essentially compared the intensity of program effort across localities with the corresponding interarea variation in program outcomes. A fundamental problem in program evaluation is that the coverage of programs and the timing of program initiatives—program placement—is not likely to be random. This is true to the extent that governmental decision rules are responsive to attributes of the targeted populations that are not measured in the data. Simple measured associations between programs and program outcomes, anticipated or unanticipated, will therefore not provide correct

Mark M. Pitt is with the Department of Economics at Brown University, Mark R. Rosenzweig is with the Department of Economics at the University of Pennsylvania, and Donna M. Gibbons is with the Department of Economics at Carleton College.

estimates of program effects. A research methodology and data base that can accommodate the existence of unobserved, location-specific attributes that influence both program placement and program outcomes are needed. This article uses Indonesia's uniquely rich data base to employ methods of analysis that reveal both the patterns of public program placement and the consequences of the programs, even if the programs are endogenously allocated.

In any country, at a point in time, program efforts vary widely across areas, even if the programs are funded and controlled by the central government. Given the limited resource capacities of the central public agency, program allocations must be rationed. The placement of programs is thus likely to depend on the expected location-specific returns to the program, which will vary across areas according to, among other attributes, their physical and demographic characteristics or endowments. If program placement is attentive to locational endowments and such endowments also influence outcomes of interest to policymakers, it is important in evaluating policies or programs to have information on endowments. It is inevitable, however, that not all exogenous locational characteristics are measured or are measurable.

Data on the spatial distribution of programs and population characteristics at more than one point in time can be used to identify program effects and the "rules" by which programs are allocated. When program placement depends on unmeasured time-persistent or permanent characteristics of locations but varies as a function of aggregate economywide trends or shocks (economic or political), cross-sectional data cannot readily be used to identify either program allocation rules or their consequences unless the strong assumption is made that some area-specific characteristics affect program placement and not, net of the programs, program outcomes. However, estimates of program effects free from the contamination of area heterogeneity bias can be obtained from estimates of how changes in local programs affect changes in local population characteristics (fixed effects).

The placement of any particular type of program is likely to be sensitive not only to the demographic characteristics of regions but also to the regional distribution of programs that are already in place. A primary goal of the placement of a program in a specific locality is to enhance access to the program. Because fees charged by government programs are nominal or zero, "access" represents the cost of traveling to a program—its distance from a spatially defined population. Gertler and van der Gaag (1990) have shown that in Côte d'Ivoire the market for medical care is rationed by the time costs involved in obtaining care from providers. If there are already similar or identical programs nearby, the initiation of a new program of the same type has a lower return compared with placing it where there are no programs of a similar type nearby. Where medical providers are more densely distributed over a fixed area, the incremental reduction in the average time cost falls. Thus, the effectiveness of a local program depends on its proximity to other programs.

The returns to a particular program may also be enhanced by the attributes of affected households. For example, the payoff to programs providing medical care may be enhanced by higher levels of education. Comprehensive information on programs is thus required, or the estimated average effects of any specific program will be biased because of omitted, correlated (program) variables. If, for example, the researcher omits variables reflecting the availability or levels of schooling when evaluating the effects of a health program whose payoff depends on the education levels of program clients, the evaluation will tend to overstate the effectiveness of the health program if health program placement is positively correlated with schooling availability. Thus, useful studies of the impact of programs must take into account the endogenous placement of programs and must use information on the proximity of as many programs as possible to the relevant individuals and households. And, of course, appropriate data on outcomes of interest to policymakers must also be used.

Existing studies and data bases have several deficiencies. Only one study (Rosenzweig and Wolpin 1986) has examined the problem of the endogeneity of program placement. That study, which used longitudinal data on nutritional status, found that inattention to this problem led to severe biases in the estimates of the effectiveness of the two programs studied (health and family planning programs). In particular, because the government evidently placed these programs first in less healthy areas, standard (cross-sectional) estimation procedures led to the erroneous inference that exposure to the programs reduced nutritional status. In fact, their estimation results indicated that the programs enhanced nutritional status once the endogeneity of the placement was "controlled." The Rosenzweig-Wolpin study thus demonstrates empirically the importance of the nonrandomness of the spatial distribution of government programs. However, the study used information on households in only 20 barrios in the Philippines and did not have comprehensive data on programs and exogenous population characteristics. Whether generalizations can be made from the study is not clear.

Existing micro data bases are not well suited for program analysis. Although many contain the necessary detail on outcomes (such as health, productivity, and education) and on relevant demographic and socioeconomic variables, they often do not have information on access to programs. And often they do not cover a sufficient number of localities to support reliable estimates of program effects. In recent years economists have merged area-specific information on programs with large household-level data sets providing location-of-residence information. In all of these cases, however, the program data are highly aggregated, so that the proximity of the households to the programs is not very well measured. For example, household data have been combined with district-level information on programs in Colombia by Rosenzweig and Schultz (1982), in Indonesia by Pitt and Rosenzweig (1985), and in Côte d'Ivoire by Strauss (1990).

In recent years new data collection initiatives (including the Living Standards Measurement Surveys) have included information on program proximity in survey instruments. Although initial surveys of this type collected data only on the distance to programs actually used by the sample respondents, the newer efforts have collected data on program proximity independent of use. However, the cross-sectional or closely spaced panel data that result from these surveys cannot be used to correct for the problem of endogenous program placement.[1]

This article reports on research based on data consisting of newly merged Indonesian household-level, cross-sectional census data and comprehensive *kecamatan* (subdistrict) -level information on programs from two time periods. These data are used to assess the effects of a variety of programs on the schooling of children by gender, child mortality, and fertility.

Section I sets out a framework to estimate both the effects of programs when program placement is nonrandom and the determinants of program placement. Section II describes the creation of the data set used in the estimation. The existing data base of Indonesia, when appropriately assembled, aggregated, and merged, offers a unique opportunity to study the determinants and consequences of program placement at a very disaggregated spatial level by using fixed-effects methods. It also allows exploration of nonlinearities in program effects. Section III presents estimates of the effects of programs and parental schooling on six outcome measures: attendance rates, by gender, of children 10–14 and 15–18 years of age; the children ever born to all women ages 25–29; and the cumulative mortality rates of children of women ages 25–29.

The analysis indicates that the proximity of grade schools, middle schools, and health programs significantly affects the school attendance of teenagers. And there is some evidence that the grade school effects are stronger among households in which mothers are less educated. There is no evidence, however, that family planning programs significantly affect any of the outcomes studied. In addition the contrast between the cross-sectional effects and results obtained using fixed effects is quite marked. The cross-sectional estimates underestimate by 100 percent the effect of grade school proximity on the schooling attendance of both boys and girls ages 10–14. The cross-sectional estimates also suggest that family planning programs increase fertility, the result being significant at the 0.05 level, whereas the results based on the fixed-effects method suggest that family planning clinics reduce fertility (although the coefficient is imprecisely measured). Estimates of the determinants of program placement confirm that they are not random with respect to unobserved factors determining outcomes and behaviors and that during 1980–86 program coverage across subdistricts was being equalized for all programs considered except one.

## I. THE ANALYTIC FRAMEWORK

This section presents the analytic framework for estimating endogenous program effects and government placement rules.

1. The problem with closely spaced panel data is that program change is likely to be small for periods shorter than two or three years. As a result, it is difficult to statistically identify program effects with any precision.

*Estimating Endogenous Program Effects*

Equation 1 is a representation of a set of program evaluation equations based on data describing programs, population characteristics, and outcomes across geographic regions that is typical in program evaluation studies that do not focus solely on one program.

$$(1)\quad H_{rit} = \Sigma_k \beta_{rk} P_{kit} + \Sigma_m \theta_{rm} Z_{mit} + \Sigma_n \Phi_{rn} E_{ni} + \mu_{ri} + \epsilon_{rit}, \qquad r = 1, \ldots, R$$

where $H_{rit}$ is policy outcome $r$ in geographic region $i$ at time period $t$ (for example, the fraction of children of a specific age and sex who are in school); $P_{kit}$ are the set of $N$ programs in the regions at time $t$; $Z_{mit}$ are the relevant socioeconomic characteristics (age, sex, level of education, and so on) of the individuals or households in the region; $E_{ni}$ are measured environmental characteristics of the region (for example, altitude, propensity to drought, and flood); $\mu_{ri}$ is a time-invariant, outcome-specific, unmeasured attribute (latent policy outcome) of the region; and $\epsilon_{rit}$ is a random, time-varying error. The $\beta_{rk}$, $\theta_{rm}$, $\Phi_{rn}$ are parameters to be estimated; $\beta_{rk}$ in particular are the estimates of the program effects on the outcomes.

Program effects may differ across households and individuals; for example, the more educated may benefit more or less from particular programs, or there may be different effects for males and females. Allowing program effects in the program evaluation equation to differ with the household attributes, $Z_{mit}$, results in the specification

$$(2)\quad H_{rit} = \Sigma_k \beta_{rk} P_{kit} + \Sigma_m \theta_{rm} Z_{mit} + \Sigma_n \Phi_{rn} E_{ni} + \Sigma_k \Sigma_m \delta_{rkm} P_{kit} Z_{mit} + \mu_{ri} + \epsilon_{rit}, \qquad r = 1, \ldots, R$$

where $\delta_{rkm}$ are the parameters describing the attribute-specific program effects. These terms provide estimates of how the effectiveness or effects of each of the programs on each outcome, $H_r$, depend on observable attributes of households. For example, are health clinics more effective at reducing child mortality for mothers who are better schooled? Are the returns to increasing the coverage of schools higher if the schools are located in areas in which schooling of parents is low? The answers to these questions are useful for allocating programs efficiently and for designing and evaluating patterns of program allocation that are intended to target benefits to certain classes of households defined by their characteristics, $Z_{mit}$.

The principal problem in obtaining estimates of the matrix of program effects described in equations 1 and 2 is that the programs may not be orthogonal to the unmeasured attributes of the localities, $\mu_{ri}$. If, for example, the government, because of financial constraints, cannot provide program support across all areas at one time, it may implement a plan for a phased distribution of programs to be allocated to the regions over time. The existence of a program in a region at any point in time is likely to be a function of the permanent latent outcomes of the region that are unobserved by the researcher ($\mu_{ri}$). Thus, for program $P_{kit}$,

$$(3)\quad P_{kit} = \Sigma_n \gamma_{1kn} E_{ni} + \Sigma_r \gamma_{2kr} \mu_{ri} + u_{kit}, \qquad k = 1, \ldots, N$$

where $\gamma_{1kn}$ and $\gamma_{2kr}$ are unknown estimable parameters characterizing the government program placement rule and $u_{kit}$ is a random time-varying error. A more dynamic representation of the governmental decision rule is one in which the coverage of programs in a locality at time $t$ influences the subsequent growth of program coverage across areas. Thus,

$$(4) \qquad P_{kit+1} - P_{kit} = \Sigma_n \alpha_{1kn} E_{ni} + \Sigma_o \alpha_{2ko} P_{oit} + \Sigma_r \alpha_{3kr} \mu_{ri} + e_{kit}, \qquad k = 1, \ldots, N$$

where $P_{oit}$ are the existing set of programs in region $i$ at time $t$ including $P_{kit}$; the $\alpha_{1kn}$, $\alpha_{2ko}$, and $\alpha_{3kr}$ are estimable parameters; and $e_{kit}$ is a random time-varying error.

Equations 3 and 4 indicate that as long as $\gamma_{2kr} \neq 0$ and $\alpha_{3kr} \neq 0$—that is, that program placement is attentive to area attributes not measured in the data—use of least squares applied to cross-sectional data to estimate program effects, from equations 1 or 2, will be biased. That is, both the region-specific outcomes and the programs are correlated with $\mu_{ri}$. One method of eliminating the bias is to eliminate the $\mu$'s from the equations, because they are the source of the correlation between the least squares residual and the regressors. With information on program placement and outcomes at two points in time for the same region, for example, a fixed-effects procedure can be implemented that sweeps out the unobservable, as in equation 5:

$$(5) \qquad H_{rit+1} - H_{rit} = \Sigma_k \beta_{rk} (P_{kit+1} - P_{kit}) + \Sigma_m \theta_{rm} (Z_{mit+1} - Z_{mit}) + \epsilon_{rit+1} - \epsilon_{rit}, \qquad r = 1, \ldots, R.$$

Thus, by relating the changes over time in outcome variables to changes overtime in program placement, the biases resulting from the endogeneity of program placement are eliminated as long as the region-specific, time-varying shocks affecting program placement in equation 3 or 4 are uncorrelated with the region-specific, time-varying disturbances in the outcome equations.

The outcome equations (1 and 2) assume that programs affect only outcomes contemporaneously. They assume that the history of programs in an area does not matter for current-period program outcomes. This assumption may be reasonable for some outcomes, such as period-specific birthrates for young women, infant mortality rates, or school enrollment rates for children of primary-school age.[2] However, the assumption that the existence of programs in the past does not influence such contemporaneous outcomes as the schooling of older children, the cumulative fertility of older women, or the health of older children is not realistic. Whether primary schools were present in an area in the past, for example, clearly affects whether or not children currently of secondary-school age will attend secondary schools, as will the existence of secondary schools in the area in the current period. And whether or not family planning programs

2. This assumes that the history of program placement is not used by households or individuals to form expectations about future program placements.

were present in the past clearly affects the current birth decisions of older women, because it will have affected their cumulative fertility.

Thus, for some outcomes, a more appropriate specification of the set of outcome equations, ignoring for simplicity differences in program effects across socioeconomic groups, is

$$(6)\qquad H_{rit} = \Sigma_j \Sigma_k \beta_{rkj} P_{kit-j} + \Sigma_m \theta_{rm} Z_{mit} + \Sigma_n \Phi_{rn} E_{ni} + \mu_{ri} + \epsilon_{rit},\quad r = 1, \ldots, R; j = 0, \ldots, J$$

where for simplicity as well we assume that there are no lagged effects of socioeconomic variables. Equation 6 differentiates program effects, $\beta_{rkj}$, by their lags of length $j$. Estimation of equation 6 using least squares also results in biased estimates of all of these program effects, for the same reason as it does for equations 1 and 2, because the programs are correlated with the latent outcome variables, as in equation 3.

With only two ($N$) time-specific observations on outcomes and programs and no retrospective information on the history of programs by region, it is obvious that not all of the lagged program effects can be estimated without bias or even at all. Indeed, no lagged program effects for lags greater than two ($N$) can be estimated at all in that case. However, if the program allocation rule is described by equation 3, differencing across the two periods can yield unbiased estimates of the contemporaneous program effects, $\beta_{rk0}$, even if such programs are endogenous and even if there is no information on lagged programs. To see this, substitute equation 3 for all of the relevant lagged programs in equation 6 so that the outcomes are functions only of the contemporaneous programs, the time-invariant latent variables, and the lagged program shocks,

$$(7)\qquad H_{rit} = \Sigma_k \beta_{rk} P_{kit} + \Sigma_k \Sigma_j \beta_{rkj} \Sigma_n \gamma_{2kn} \mu_{ni} + \Sigma_k \Sigma_j \beta_{rkj} u_{kit-j} + \Sigma_m \theta_{rm} Z_{mit} + \Sigma_n \Phi_{rn} E_{ni} + \mu_{ri} + \epsilon_{rit},\quad r = 1, \ldots, R; j = 1, \ldots, J$$

so that

$$(8)\qquad H_{rit} = \Sigma_k \beta_{rk0} P_{kit} + \Sigma_m \theta_{rm} Z_{mit} + \Sigma_n \Phi_{rn} E_{ni} + \mu^*_{ri} + \epsilon^*_{rit},\quad r = 1, \ldots, R.$$

Equation 8 is similar to equation 1 except that the fixed effect $\mu^*_{ri}$ contains the lagged program effects and program responses to the set of area outcome-specific endowments as well as the endowment specific to the outcome $r$.

Differencing equation 7 across two periods thus yields

$$(9)\qquad H_{rit+1} - H_{rit} = \Sigma_k \beta_{rk0} (P_{kit+1} - P_{kit}) + \Sigma_k \Sigma_j \beta_{rkj} (u_{kit+1-j} - u_{kit-j}) + \Sigma_m \theta_{rm} (Z_{mit+1} - Z_{mit}) + \epsilon_{rit+1} - \epsilon_{rit},\quad r = 1, \ldots, R$$

which provides an unbiased estimate of $\beta_{rk0}$ as long as program changes do not respond to lagged program shocks. Thus, no matter how many lags there are in program effects (equation 6), it is still possible to obtain unbiased estimates of the contemporaneous effects of the programs with only two sets of period-specific observations. However, these contemporaneous effects are not the full effects of the programs if there are lags, because the current programs will have effects on future outcomes that cannot be estimated. Moreover, from equation 8, if the program allocation rules conform to equation 4 and are dynamic, then differencing across the two periods does not yield unbiased estimates of even contemporaneous program effects when there are important lagged program effects, because the history of program shocks influences the current program allocations.

### *Estimating Government Placement Rules*

The effects of endogenously placed programs on outcomes of policy interest thus can be estimated under a plausible set of assumptions even with two sets of cross-sectional data by using fixed-effects techniques to estimate the sets of equations described by equations 1 and 2 or equation 6. Because such estimates can also be used to obtain estimates of the fixed effects themselves (for each outcome), the program placement rules described by equations 3 and 4 can also be estimated. That is, the analysis can be used to assess whether the spatial distribution of programs tends to equalize spatial differences in outcomes (such as health) or to exacerbate them. For example, it can be ascertained whether localities tending to exhibit high child mortality rates, net of programs, are more likely to have received health programs or whether areas of high fertility are more likely to have received family planning programs.

Theory does provide much guidance in predicting how public programs are allocated across population groups. Altruism theories of public behavior suggest that the government would allocate more programs to those areas in which latent outcomes (such as health and schooling) were least. Thus, areas with high fertility would receive the greatest coverage of family planning clinics, and areas with high child mortality would receive the greatest coverage of health clinics. In this model the government allocates programs to compensate populations that are poorly endowed with latent outcomes. Alternatively, pressure-group theories suggest that the government may respond to those who have high demands for these outcomes by providing them with a disproportionate share of program resources. Households having the highest latent outcomes would derive the greatest private benefit from these programs and may be willing to lobby hardest for these resources. Efficiency criteria and externalities such as the transmission of disease, as in the model of Rosenzweig and Wolpin (1986), may also influence the allocation of programs. Program allocations may also respond to differentials in gross returns to programs across population groups—defined by parental schooling, for example—resulting from nonlinearities in the program effects function (as in equation 2).

An econometric problem in estimating the program placement equations is that the estimates of the fixed effects contain error, which will lead to biases

(Pitt, Rosenzweig, and Hassan 1990). However, with two estimates of each latent outcome based on two matched cross-sections, the set of estimated endowments from one cross-section can be used as instruments for the set of estimated endowments from the other as long as the time-varying or transitory errors ($\epsilon_{rit}$) are independent and identically distributed, as is necessary to assume if there are lagged program effects. Thus, the parameters of equations 3 and 4 can be estimated by using instrumental variable methods to correct for "errors in variables" in the measurement of $\mu_{ri}$, with noncontemporaneous estimates of fixed effects as identifying instruments.

The dynamic program allocation equation (4) has the coverage of programs, $P_{kit}$, on both sides of the equation, suggesting an additional errors-in-variables problem in its estimation. If $P_{kit}$ is measured with error, own-program coverage will have a possibly spurious negative effect on program growth, and the estimated effects of other programs, $P_{mit}$ $(m \neq k)$, on the growth of coverage of program $k$ will be biased toward zero. Program coverage in a period before year $t$ can be used as instruments for the set of (level) program coverage in period $t$.[3]

## II. The Data Set

To create a data set that provides answers to questions about the effects of programs, the influence of nonrandom program placement on program assessments based on cross-sectional data, and the relation between area endowments and program allocations, information is needed at, minimally, two points in time on programs, outcomes of program effects, and characteristics of geographical areas that may have influenced program placement. The data used in the empirical research combine the 1986 and 1980 Potensi Desa of Indonesia (PODES), the 1976–77 Fasilitas Desa of Indonesia (FASDES), the 1985 Intercensal Population Survey of Indonesia (SUPAS), and the 1980 Population Census of Indonesia. All the surveys were carried out by the Central Bureau of Statistics of Indonesia (Biro Pusat Statistik [BPS]).

The 1986 and 1980 PODES and the 1976–77 FASDES provide information at the village level on government programs such as schools, family planning clinics, health centers, and sources of water for drinking and bathing. The surveys also include data on population; on other infrastructure, such as marketplaces, banks, factories, types of roads, recreation facilities, communication facilities, and electricity; and on area-specific geophysical characteristics, including altitude, land type, proximity to coastline, and the history of natural disasters. Approximately 67,000, 62,000, and 58,000 villages were surveyed in the 1986 PODES, the 1980 PODES, and the 1976–77 FASDES, respectively.

3. If program coverage is measured with error, classical errors-in-variables bias will contaminate the estimates of the program effects from equation 1. Parameters will be biased toward zero. Unfortunately, it is difficult to find instruments that can be used to correct for this potential bias. To our knowledge, no other study of program effects has addressed the issue of measurement error in the program coverage variables.

The 1980 Population Census of Indonesia and the 1985 SUPAS provide information, including census block information that can be mapped into village of residence, for a stratified random sample of households throughout Indonesia on individual schooling, labor force participation, marriage, fertility, birth control, and child mortality. They also provide data on selected household assets, structure, land area, and cooking and bathing facilities. The detailed questionnaire for the 1980 census was used for a stratified random sample of 1,502,075 households containing 7,234,634 individuals, approximately 5 percent of the total population of Indonesia. The 1985 SUPAS surveyed 126,370 households across Indonesia containing 602,885 individuals. An important shortcoming of the survey data is that information on income, expenditures, and the total value of household assets is not available.[4]

To obtain a data set that combines the program and household information at two points in time (1980 and 1985), it is necessary to link all of the data sets by geographical area. Indeed, the assumption that program placement rules depend on regional-level characteristics means that data over time must be linked at that level. Thus, it is not necessary to have household-level longitudinal data to appropriately estimate program effects.[5] Because the data on individual Indonesian households are not longitudinal, the successive cross-sections are aggregated to the kecamatan level and matched at that level. Because the underlying data are individual, however, the specification need not be restricted to linear forms with respect to individual or household characteristics. Thus, for example, logarithmic specifications can be tried, aggregating up from log transforms of the micro variables.

Because information at the village level provides the most accurate information on households' proximity to programs, it would have been desirable to link the data at that level. Unfortunately, this was not possible because there were administrative changes in the boundaries and names of villages throughout 1976–86. Also, the BPS updated geographical location codes at the village (*desa*), subdistrict (*kecamatan*), district (*kabupaten*), and province (*provinsi*) levels three times from 1976 through 1986: in 1980 coinciding with the Population Census (Sensus Penduduk), in 1983 coinciding with the Agricultural Census (Sensus Pertanian), and in 1986 coinciding with the Economic Census (Sensus Ekonomi). In each of these years a PODES survey was also conducted as part of the corresponding census. Thus, all of the original data sets used in this research have different location codes. The BPS does not comprehensively document location code changes over time. However, for 1980–86 the BPS provides a master list with names and codes of all villages, subdistricts, districts, and provinces for the years of the updating. These lists, called "master files," do not enable data sets to be matched accurately and consistently at the village level but do enable the codes to be tracked over time

4. The implications of the absence of this information are discussed below.

5. This assumes that program placement rules are attentive only to fixed geographic characteristics.

at the kecamatan level. The appendix provides details on the matching procedures that were used.

This article examines the effects of, and placement rules for, three types of programs: schooling institutions by level (grade, middle, and high school); health clinics, especially the *puskesmas* program; and family planning. Table 1 provides information on the coverage of these programs in Indonesia based on the aggregated and matched PODES data for 1980 and 1986. The proportion of households living in a village with each of the programs increased considerably during that period. For example, 93 percent of households resided in a village with a grade school in 1986, up from 76 percent in 1980. In 1986 42 percent of households resided in a village with a government health clinic, compared with 24 percent in 1980. Given that the methodology exploits the change in program coverage over time, this evident substantial change is a useful feature of the constructed data set.

The spatial correlations among the program variables indicate that kecamatans that have a high degree of exposure to one type of program are likely also to have relatively high coverage of the other programs—all spatial correlations of program variables are positive and statistically significant. For example, in 1980 the correlation between the coverage of middle schools and high schools is 0.72 and that between grade schools and family planning programs is 0.48. Thus, to the extent that each program has cross-effects—that is, each program affects outcomes in addition to those it is intended to influence—it may be important in the Indonesia context to estimate program effects jointly to appropriately evaluate the effects of any one program type on any particular outcome (Rosenzweig and Wolpin 1982). For example, family planning programs may reduce not only births but also child mortality rates. If the effects of health clinics on child mortality are estimated without considering the presence of family planning clinics, given that both tend to be in the same areas, the effects of the health clinics may be overestimated. We present estimates below of the extent to which the changes in program coverage during 1980–86 have reduced or increased the spatial correlation in programs.

The 1985 SUPAS and 1980 Population Census data sets include sample weighting factors (frequency weights) for both individuals and households. In both

Table 1. *Institutional and Program Coverage in Indonesia, 1980 and 1986*
(percentage of households residing in a village with a program or institution)

| *Institution or program* | *1980* | *1986* | *Growth (percent)* |
|---|---|---|---|
| Grade school | 74.4 | 93.0 | 25.0 |
| Middle school | 26.7 | 39.0 | 46.1 |
| High school | 10.3 | 17.7 | 71.8 |
| Family planning clinic | 45.9 | 76.5 | 66.7 |
| Health clinic (*puskesmas*) | 24.4 | 42.4 | 73.8 |

*Note:* Data are based on 3,302 matched *kecamatans*.
*Source:* The 1980 and 1986 PODES data set from the government of Indonesia.

data sets separate records describe the household characteristics and each individual's characteristics. The weighted variables are summed to aggregate these individual and household data at the subdistrict level. (A weighted variable is the original sample variable multiplied by the appropriate sample weight.) The resulting aggregated data thus are representative of the population of Indonesia in both 1980 and 1985. There were 3,179 aggregated kecamatan "observations" for the 1985 SUPAS and 3,253 for the 1980 Population Census.

There are two principal advantages to aggregating from micro data of the type available in the Census and SUPAS samples. First, there can be an appropriate matching of dependent and independent variables. For example, if the dependent variable characterizes children of a given age, then it is possible to obtain information on the parents of those children. Second, as noted, the appropriate aggregation of any micro functional form can be obtained. Any nonlinearities hypothesized about relations between independent and dependent variables at the household or individual level can be appropriately aggregated.

We create three types of "outcome" variables characterizing the schooling of children, fertility, and child mortality. For schooling, there are four attendance rates for female and male children ages 10–14 and 15–18. Fertility is measured by the number of children ever born to all women ages 25–29. And child mortality is measured by the cumulative child death rates for all women ages 25–29. Selecting the vital rates of relatively young women minimizes the potential influence of lagged program effects; however, such lagged effects may be important for the estimates of the attendance rates of the children ages 15–18.

The independent variables for the school attendance outcomes are the mean schooling attainment and age for mothers of children ages 10–14 and 15–18, and the mean schooling attainment for heads of the households in which children ages 10–14 and 15–18 reside. The independent variable for the fertility and child mortality outcomes is the average schooling attainment for women ages 25–29. Table 2 provides the means of the outcome and independent variables for 1980 and 1985 and the number of matched kecamatans for each variable. These figures indicate that, just as program exposure for the basic schooling, health, and fertility programs has increased significantly from 1980 through 1986, the basic human capital outcome indicators—school attendance (particularly for females) and child survival—have also increased substantially during the period, although fertility has dropped only marginally. Of course, the correlation in the overall trends in program exposure and the outcomes cannot be used to infer much about the effectiveness of the programs. For example, it is clear from table 2 that the average schooling levels of parents (and potential parents) has also increased during the period. Here the basic methodology for discerning the impact of programs essentially tests whether the changes in outcomes were greater in areas in which there were greater changes in program population (village) coverage net of changes in the schooling levels of parents.

Table 2. *Policy Outcomes and Independent Variables for Matched* Kecamatans *in Indonesia, 1980 and 1985*

| *Variable* | *Mean* 1980 | *Mean* 1985 | *Number of matched kecamatans* |
|---|---|---|---|
| *Policy outcome* | | | |
| School attendance(percentage of total in age group) | | | |
| Females ages 10–14 | 76.9 | 88.1 | 3,043 |
| | (14.1) | (16.3) | |
| Males ages 10–14 | 81.6 | 90.2 | 3,048 |
| | (12.5) | (14.8) | |
| Females ages 15–18 | 29.2 | 44.8 | 2,887 |
| | (19.7) | (33.8) | |
| Males ages 15–18 | 39.2 | 52.9 | 2,548 |
| | (20.3) | (33.0) | |
| Number of children ever born to women ages 25–29 | 2.60 | 2.45 | 3,014 |
| | (0.69) | (0.79) | |
| Cumulative mortality rate[a] of children of women ages 25–29 | 14.2 | 8.00 | 3,005 |
| | (8.60) | (10.3) | |
| *Independent variable* | | | |
| Schooling attainment (years) | | | |
| Mothers of children ages 10–14 | 2.65 | 3.54 | 3,043 |
| | (1.52) | (2.05) | |
| Household heads in households with children ages 10–14 | 3.61 | 4.53 | 3,043 |
| | (1.62) | (2.62) | |
| Mothers of children ages 15–18 | 2.08 | 3.06 | 2,887 |
| | (1.46) | (2.21) | |
| Household heads in households with children ages 15–18 | 3.11 | 4.17 | 2,887 |
| | (1.53) | (2.46) | |
| Women ages 25–29 | 4.20 | 4.78 | 3,014 |
| | (1.86) | (2.16) | |

*Note:* Standard deviations are in parentheses.
a. Percentage of total live births that died.
*Source:* The 1980 census and the 1985 SUPAS.

To test whether program effects on the school attendance, child mortality, and fertility outcomes differ by the schooling level of women, the micro sample is split into three education groups—women (for the relevant school-age children or in the relevant age group) with no schooling, women with one to five years of schooling, and women with six or more years of schooling.[6] The kecamatan aggregation was performed for each of these three groups.

## III. Estimates of Program Effects

Tables 3, 4, and 5 report the estimates of program and parental schooling effects on the six outcome measures. In each table, three sets of estimates are reported. All the regression estimates reported here provide *t*-ratios based on

6. The 1985 SUPAS indicates that 19 percent of women ages 20–40 had received no schooling. Thirty-five percent had one to five years of schooling.

Table 3. *Estimates of the Determinants of School Attendance for Females and Males Ages 10–14*

| | Females | | | Males | | |
|---|---|---|---|---|---|---|
| | *Weighted least squares* | | | *Weighted least squares* | | |
| *Variable* | *Excluding health and family planning clinics* | *Including health and family planning clinics* | *Fixed-effects methodology* | *Excluding health and family planning clinics* | *Including health and family planning clinics* | *Fixed-effects methodology* |
| Schooling attainment, mothers | 0.035 | 0.035 | 0.034 | 0.0017 | 0.0033 | 0.010 |
| | (10.5) | (10.7) | (3.51) | (0.62) | (1.20) | (3.65) |
| Schooling attainment, household heads | 0.024 | 0.023 | 0.0075 | 0.039 | 0.037 | 0.0078 |
| | (6.37) | (6.17) | (2.19) | (12.8) | (12.3) | (3.05) |
| Age of mother | 0.010 | 0.0094 | −0.0030 | 0.0071 | 0.0059 | −0.0013 |
| | (8.31) | (7.35) | (2.36) | (6.59) | (5.43) | (1.37) |
| Owned land per household ($\times 10^{-4}$) | 0.19 | −0.20 | 1.6 | −1.1 | −1.2 | 1.4 |
| | (0.34) | (0.35) | (2.12) | (2.25) | (2.50) | (2.20) |
| Proportion of households in urban areas | −0.065 | −0.062 | −0.029 | −0.042 | −0.040 | 0.012 |
| | (7.15) | (6.81) | (0.57) | (5.62) | (5.15) | (0.25) |
| Proportion of households in villages | | | | | | |
| With grade school | 0.039 | 0.048 | 0.11 | 0.052 | 0.041 | 0.088 |
| | (3.81) | (4.68) | (5.90) | (2.97) | (4.79) | (5.61) |
| With middle school | 0.067 | 0.085 | 0.085 | 0.050 | 0.073 | 0.054 |
| | (5.82) | (6.88) | (3.87) | (5.45) | (7.45) | (3.02) |
| With high school | 0.011 | −0.14 | −0.035 | 0.0010 | 0.0003 | −0.011 |
| | (0.88) | (1.13) | (1.57) | (0.09) | (0.03) | (0.55) |
| With health clinic (*puskesmas*) | — | −0.028 | 0.033 | — | −0.031 | 0.001 |
| | | (3.19) | (2.01) | | (3.92) | (0.07) |
| With family planning clinic | — | −0.015 | 0.0076 | — | −0.029 | 0.010 |
| | | (2.25) | (0.85) | | (5.59) | (1.26) |
| Time trend | — | — | 0.052 | — | — | 0.037 |
| | | | (9.43) | | | (7.49) |
| Constant | 0.15 | 0.19 | — | 0.38 | 0.43 | — |
| | (3.01) | (3.68) | | (8.54) | (9.58) | |
| $R^2$ | 0.53 | 0.53 | — | 0.48 | 0.49 | — |
| Number of *kecamatans* | 2,904 | 2,904 | 2,874 | 2,904 | 2,904 | 2,881 |

*Note:* Weights are the sample number of households. The least squares estimates are based on the 1980 merged cross-sectional subdistrict data. The fixed-effects methodology uses the 1980 and 1985 data sets. *Kecamatans* with fewer than 10 survey households were excluded from the analysis. Absolute values of asymptotic $t$-ratios are in parentheses.

Huber's method for calculating the parameter covariance matrix (Huber 1967; see also White 1980). The estimates of the $t$-ratios are generally consistent even if there is heteroskedasticity or clustered sampling or weighting not correctly accounted for in the aggregation of the micro data or in the weighting of the least squares estimates. The first set is obtained using weighted least squares, where the weights are the sample number of households, based on the 1980 merged cross-sectional, kecamatan data. The first set includes in the specification only the program variables directly relevant to the outcome.[7] Thus, for the school attendance measures, the health and family planning clinics are excluded. For the fertility outcome, the three school types and the health clinic are excluded. For the child mortality outcome, the family planning clinic and school programs are excluded. This first specification corresponds to that which is prevalent in the evaluation research literature, where estimates of program effects are based on cross-sectional data and tend to focus on a narrow set of programs that are assumed to be directly relevant to the outcome being studied. The second set of estimates differs from the first only in that all of the program measures are included. The third specification implements the fixed-effects methodology, using both the 1985 and 1980 data sets and including all of the program measures.[8]

All of the specifications also include the amount of land owned by each household and the proportion of households residing in urban areas. The fixed-effects estimates also allow for a time trend in all dependent variables. Thus, those estimates of program effects are net of aggregate trends in the outcomes that are visible in table 2. Although the fixed-effects estimates are net of the influence of both locality-specific fixed effects and of aggregate time trends, such estimates based on only two points in time cannot control for the potential influence of area-specific time trends. This may be particularly important for the estimates, given the absence of information on incomes. Thus, significant variation in economic growth rates across areas might cause bias in the fixed-effects estimates.[9]

Table 3 reports the estimates of the determinants of school attendance for females and males ages 10–14. From these, we can draw two conclusions, which are applicable to both gender groups: first, in the cross-section, exclusion of the evidently relevant health and family planning clinics from the specification results in underestimates by up to 50 percent of the impacts of the presence of both grade schools and middle schools on the school attendance of 10- to 14-year-

7. Kecamatans with fewer than ten survey households were excluded from the analysis.

8. We have also estimated, using random effects methods, the same specifications based on the pooled 1980 and 1985 data. The results from these estimates are not very different from those obtained from the cross-sectional 1980 data. We show the cross-sectional 1980 estimates as these correspond to those obtained in the literature. If data analysts had pooled time-series, cross-sectional data of the type we have constructed, with both program proximity information and relevant outcomes, fixed-effects methods would likely be employed.

9. For example, high-growth areas may have sharply increasing demands for additional schooling for their children and for additional schools, to which the government may respond.

olds. Second, use of the cross-sectional data, not taking into account the possibly nonrandom spatial location of programs, results in an underestimate by 100 percent of the effect of being proximate to a grade school on the school attendance of both males and females ages 10–14.

The cross-sectional estimates indicate that an increase in the coverage of villages with grade schools to 100 percent, from, say, the 1980 figure of 74 percent, would increase school attendance by 1.0 to 1.2 percentage points, whereas the fixed-effects estimates indicate that the increase would be by 2.2 to 2.8 percentage points. The fixed-effects estimates indicate that an increase to universal coverage of grade schools would have raised the school attendance rates, based on the 1980 census figures in table 2, to 79.7 and 83.7 percent for females and males, respectively. The estimates also indicate that a similar increase by about 25 percentage points in the coverage of middle schools would have raised school attendance rates by an additional 2.1 percentage points for females and by an additional 1.4 percentage points for males. This is substantially less than the growth in the rates that occured during 1980–85, when coverage of grade schools increased to 93 percent and coverage of middle schools increased by only 19 percentage points (a substantial relative increase). The best estimates thus indicate that the growth in the spatial coverage of grade schools and middle schools played a relatively small, but not insignificant, part in the growth in school attendance of females and males in the 10–14 age group during this period.

The cross-sectional estimates of the effects of the health and family planning clinics on school attendance are not credible, being both negative and statistically significant, whereas the fixed-effects estimates indicate that both programs may have a positive impact on school attendance. The positive effect of the health clinic on the school attendance of females is statistically significant, just as it is for the 10–14 age group. The cross-sectional estimates of the effects of the schooling attainment of household heads, almost all of whom are male, on school attendance rates are also considerably higher than the fixed-effects estimates, whereas the estimates of the effects of maternal schooling are relatively robust to estimation procedure. The preferred fixed-effects estimates indicate that, for both males and females, the effect of the mother's schooling attainment on school attendance is greater than that of the male head's schooling. The differences are statistically significant for both gender groups. The fixed-effects point estimates indicate that for each one-year increase in the number of years of schooling of mothers, the school attendance rate of their female children rises by 3 percentage points and that of their male children by 1 percentage point. For each one-year increase in the schooling attainment of household heads, school attendance rises by 3/4 percentage point for both males and females. The increase in schooling attainment (by less than one year) for mothers and household heads between 1980 and 1985 thus accounts for only a small part of the actual increase in school attendance of almost 19 percentage points during that period.

The contrast between the set of cross-sectional program effects estimates and those obtained using fixed effects is even more marked for the school attendance rates of the 15- to 18-year-olds, as reported in table 4. The cross-sectional estimates indicate that in kecamatans in which there are higher proportions of the population located in villages with middle and high schools, attendance rates for this age group are significantly higher; however, the rates are significantly lower where there is a greater coverage of grade schools, health clinics, and family planning clinics. In contrast, the fixed-effects estimates indicate that grade schools are significantly positively related to the school attendance of 15- to 18-year-olds (only at the 10 percent level for females), as are the health clinics for females, but there are no effects of the coverage of either of the other school types or of the family planning clinics. The difference between the effects of the schooling attainment of the mother and household head on school attendance is also substantially reduced when the fixed effects procedure is used compared with the cross-sectional estimates. The point estimates of the grade school effects and the parental schooling variables are relatively small and indicate that the growth in grade school coverage and in the schooling of parents in Indonesia between 1980 and 1986 cannot alone account for the 53 percent increase in school attendance among female teenagers ages 15–18 or the 35 percent increase among males over that period, as exhibited in table 2.

Similar dramatic differences in inferences about program effects by estimation procedure and non-credible results from cross-sectional estimates are seen in table 5, which presents the estimates of the determinants of fertility and child mortality. For fertility, the estimates that are based on the cross-sectional association between program coverage and outcomes and that exclude the effects of alternative programs indicate that family planning programs increase fertility, with the result significant at the 0.05 level! The point estimate of the effect is not influenced very much by the inclusion of other programs in the specification, although the significance level drops. The fixed-effects estimate, however, indicates that the increased coverage of family planning clinics does reduce fertility, although the coefficient is imprecisely measured and the impact is very small.

The cross-sectional fertility results indicate that the presence of middle and high schools significantly reduces fertility, but this result also appears to be due to the nonrandomness of school placement, because these effects disappear when the fixed-effects procedure is used to obtain estimates. The latter estimates, however, indicate that the presence of grade schools positively affects fertility, but by only a small amount. Both the cross-sectional and the fixed-effects estimates indicate that the health clinics also positively affect fertility, although the effects are also small. An increase in the number of grade schools to attain universal coverage across villages would increase fertility by only 0.07 children. A doubling of the coverage of health clinics from 1980 levels would have increased fertility by a similar amount. Finally, the cross-sectional method, based on the full specification, substantially underestimates (by a factor of 36, compared with the fixed-effects method) the negative effect of maternal school-

Table 4. *Estimates of the Determinants of School Attendance for Females and Males Ages 15–18*

| Variable | Females: Weighted least squares: Excluding health and family planning clinics | Females: Weighted least squares: Including health and family planning clinics | Females: Fixed-effects methodology | Males: Weighted least squares: Excluding health and family planning clinics | Males: Weighted least squares: Including health and family planning clinics | Males: Fixed-effects methodology |
|---|---|---|---|---|---|---|
| Schooling attainment, mothers | 0.011 | 0.011 | 0.026 | 0.040 | 0.038 | 0.022 |
| | (2.47) | (2.53) | (5.14) | (8.44) | (8.09) | (4.69) |
| Schooling attainment, household heads | 0.058 | 0.057 | 0.039 | 0.090 | 0.088 | 0.036 |
| | (13.8) | (13.6) | (9.02) | (19.50) | (19.10) | (9.28) |
| Age of mother | 0.015 | 0.014 | 0.0049 | 0.016 | 0.014 | 0.0010 |
| | (12.40) | (11.2) | (3.61) | (11.2) | (9.85) | (0.75) |
| Owned land per household ($\times 10^{-4}$) | −3.1 | −2.90 | 0.570 | −0.510 | −0.500 | −1.00 |
| | (5.39) | (5.03) | (0.48) | (7.31) | (7.19) | (0.98) |
| Proportion of households in urban areas | 0.28 | 0.031 | −0.110 | 0.046 | 0.051 | 0.016 |
| | (2.15) | (2.31) | (1.34) | (3.34) | (3.58) | (0.20) |
| Proportion of households in villages | | | | | | |
| With grade school | −0.080 | −0.067 | 0.045 | −0.069 | −0.042 | 0.054 |
| | (8.28) | (6.64) | (1.76) | (5.84) | (3.45) | (2.07) |
| With middle school | 0.086 | 0.110 | 0.011 | 0.110 | 0.160 | 0.049 |
| | (6.07) | (7.54) | (0.34) | (7.24) | (9.41) | (1.53) |
| With high school | 0.130 | 0.130 | 0.0086 | 0.120 | 0.120 | 0.0058 |
| | (6.69) | (6.65) | (0.26) | (6.12) | (6.15) | (0.18) |
| With health clinic (*puskesmas*) | — | −0.057 | 0.078 | — | −0.078 | 0.034 |
| | | (5.10) | (3.26) | | (5.96) | (1.51) |
| With family planning clinic | — | −0.012 | −0.010 | — | −0.037 | 0.0034 |
| | | (1.69) | (0.69) | | (4.46) | (0.31) |
| Time trend | — | — | 0.066 | — | — | 0.062 |
| | | | (7.48) | | | (7.08) |
| Constant | −0.530 | −0.480 | — | −0.470 | −0.380 | — |
| | (10.20) | (8.99) | | (7.50) | (6.07) | |
| $R^2$ | 0.73 | 0.74 | — | 0.65 | 0.66 | — |
| Number of *kecamatans* | 2,899 | 2,899 | 2,753 | 2,903 | 2,903 | 2,805 |

*Note:* Weights are the sample number of households. The least squares estimates are based on the 1980 merged cross-sectional subdistrict data. The fixed-effects methodology uses the 1980 and 1985 data sets. *Kecamatans* with less than 10 survey households were excluded from the analysis. Absolute values of asymptotic $t$-ratios are in parentheses.

Table 5. *Estimates of the Determinants of Fertility and Child Mortality for Women and Mothers Ages 25–29*

| Variable | *Children ever born* | | | *Child mortality* | | |
|---|---|---|---|---|---|---|
| | *Weighted least squares* | | | *Weighted least squares* | | |
| | *Excluding health clinics and schools* | *Including health clinics and schools* | *Fixed-effects methodology* | *Excluding family planning clinics and schools* | *Including family planning clinics and schools* | *Fixed-effects methodology* |
| Schooling attainment, women ages 25–29 | −.015 | −0.0023 | −0.083 | −0.011 | −0.010 | −0.0095 |
| | (1.61) | (0.28) | (8.03) | (10.0) | (9.49) | (6.17) |
| Age, women, ages 25–29 | −0.10 | −0.082 | 0.27 | −0.027 | −0.022 | 0.0073 |
| | (2.20) | (1.78) | (13.6) | (4.03) | (3.89) | (1.98) |
| Owned land per household ($\times 10^{-2}$) | 0.10 | 0.12 | −0.021 | −0.015 | −0.010 | −0.0088 |
| | (3.72) | (3.62) | (0.76) | (4.81) | (2.97) | (2.30) |
| Proportion of households in urban areas | −0.29 | −.087 | −0.25 | −0.014 | −0.0027 | −0.027 |
| | (6.01) | (1.31) | (0.78) | (2.26) | (0.31) | (0.62) |
| Proportion of households in villages | | | | | | |
| With grade school | — | 0.06 | 0.26 | — | −0.0095 | −0.0016 |
| | | (1.03) | (3.72) | | (1.41) | (0.13) |
| With middle school | — | −0.19 | 0.017 | — | −0.036 | −0.0078 |
| | | (2.26) | (0.19) | | (3.92) | (0.54) |
| With high school | — | −0.33 | 0.14 | — | 0.00031 | 0.019 |
| | | (3.49) | (1.33) | | (0.03) | (1.21) |
| With health clinic (*puskesmas*) | — | 0.20 | 0.23 | 0.011 | 0.012 | 0.011 |
| | | (3.22) | (3.37) | (1.57) | (1.56) | (0.97) |
| With family planning clinic | 0.075 | 0.071 | −0.018 | — | 0.034 | −0.0035 |
| | (1.99) | (1.77) | (0.05) | | (6.82) | (0.55) |
| Time trend | — | — | −0.25 | — | — | −0.056 |
| | | | (10.6) | | | (14.9) |
| Constant | 5.28 | 4.70 | — | 0.90 | 0.77 | — |
| | (4.40) | (3.88) | | (5.97) | (5.20) | |
| $R^2$ | 0.062 | 0.074 | — | 0.13 | 0.14 | — |
| Number of *kecamatans* | 2,904 | 2,904 | 2,862 | 2,904 | 2,904 | 2,856 |

*Note:* Weights are the sample number of households. The least squares estimates are based on the 1980 merged cross-sectional subdistrict data. The fixed-effects methodology uses the 1980 and 1985 data sets. *Kecamatans* with fewer than ten survey households were excluded from the analysis. Absolute values of asymptotic $t$-ratios are in parentheses.

ing on fertility. The negative effect is statistically significant when the fixed effects are taken into account.

The cross-sectional estimates of the determinants of child mortality based on the more complete specification of programs indicates that family planning clinics lower child survival whereas middle schools increase survival. In contrast, the fixed-effects estimates indicate that program coverage of schools, health clinics, or family planning clinics does not have a significant impact on child mortality; only the schooling attainment of women appears to affect child survival, a result that is robust to the estimation method. The point estimates indicate that for each year of schooling acquired by women ages 25–29, child mortality declines 7 percent. The small (just over a half-year) increase in the schooling attainment of women ages 25–29 between 1980 and 1985 cannot account for the 43 percent decline in child mortality among women in this age group during this period.

Thus, estimated program effects are often small or insignificant even when the influence of unmeasured area endowments on outcomes and program placement are taken into account. One reason is that the effects may not be linear, as in equation 2. To test whether program effects differ by the schooling level of women, the model is reestimated using fixed effects based on the data divided into the three schooling groups of women—no schooling, one to five years of schooling, and six or more years of schooling. Using the pooled data set of all three schooling groups, all coefficients were allowed to differ by group and three tests were performed: a test of whether the program coefficients differed across all three groups and tests of whether the program effects of the lowest (zero years of schooling) and the highest (six or more years) female schooling groups differed.

The test statistics are reported for each of the dependent variables in table 6. Of the six outcomes, in only two (school attendance for both males and females ages 10–14) were there differences in program effects across the three schooling groups of women; this difference arose between the lowest and highest schooling group for both females and males. Inspection of these estimates indicates that the only significant difference in program effects was in the influence of the coverage of grade schools on attendance. The stratified estimates indicate that grade school coverage has a significantly higher positive effect on school attendance for teens ages 10–14 among the lowest educational strata of women (mothers) compared with the highest. In particular, the point estimates indicate that grade school proximity has no effect on school attendance of males ages 10–14 whose mothers have more than five years of schooling. The coefficient is just slightly above that estimated from the whole sampled population reported in table 3 (0.11 for the two lowest education groups of women compared with 0.09 in table 3). The grade school effect on school attendance of females ages 10–14 whose mothers have no schooling is twice that among the women with one or more years of schooling. The coefficient for the lowest schooling group is 0.14, compared with the overall estimate of 0.11 in table 3, and it is 0.07 for the

Table 6. *Tests of Differences in Program Effects by Schooling Class of Women*

| Outcome variable | *Test statistic for differences in coefficients* All three schooling groups | Lowest and middle schooling group | Lowest and highest schooling group |
|---|---|---|---|
| School attendance (percentage of total in age group) | | | |
| Females ages 10–14 | 1.60[a] (12,3460) | 1.00 (6,3460) | 2.47[b] (6,3460) |
| Males ages 10–14 | 2.30[b] (12,3489) | 0.92 (6,3488) | 2.82[b] (6,3489) |
| Females ages 15–18 | 1.16 (12,3155) | 0.76 (6,3155) | 1.42 (6,3155) |
| Males ages 15–18 | 1.15 (12,3263) | 0.80 (6,3263) | 1.53 (6,3263) |
| Number of children ever born to women ages 25–29 | 1.10 (12,3327) | 0.92 (6,3327) | 1.12 (6,3327) |
| Cumulative mortality rate[c] of children of women ages 25–29 | 1.51 (12,3368) | 1.57 (6,3268) | 2.54[b] (6,3268) |

*Note:* Test statistics are *F*-statistics from weighted fixed-effects estimates of program effects. The null hypothesis for each test is that there is no difference in the coefficients. The three schooling groups are for women ages 20–40. The groups are no schooling (19 percent), one to five years of schooling (35 percent), and six or more years of schooling (46 percent). Degrees of freedom are in parentheses.

a. Significance level is at least 0.10.

b. Significance level is at least 0.05.

c. Percentage of total births that died.

two highest schooling groups. No other program effects differed across the schooling groups with respect to these school attendance variables.

These results suggest that the linear specification, with respect to female schooling, is a reasonable approximation for the Indonesian population. The results also indicate that the average returns to increasing the number of grade schools may be higher if they are located in populations in which schooling levels of women are low. No other basis for targeting programs is discernible from these results with respect to the criterion of different gross program returns.

## IV. How Are Programs in Indonesia Targeted?

### *The Determinants of the Cross-Kecamatan Variation in Program Coverage*

The marked differences between the cross-sectional and fixed-effects estimates of program effects suggest that the cross-area placement of programs is significantly correlated with time-persistent unmeasured factors influencing the policy outcomes. In this section, estimates are presented of the determinants of program placement, including these latent effects. Specifically, the cross-kecamatan variation in program coverage for each program in 1980 is related to the cross-subdistrict variation in the latent or fixed outcome factors that are net of program effects, obtained from the fixed-effects estimates; measures of kecamatan endowments, such as altitude, disaster history, and geographical location; and measured characteristics of the kecamatan population, such as schooling levels.

To estimate the latent outcome variables, the analysis uses the fixed-effects coefficient estimates of tables 3, 4, and 5 and applies them to the 1980 and 1985–86 data sets. There are thus two measures of each of the six time-invariant factors for each kecamatan corresponding to the six outcome variables. Because each estimated factor contains measurement error, use of either set as regressors in the specification that determines program placement would result in bias. Instead, one set of the measures of the fixed factors (those from the 1985–86 data set) is used as instruments for the other set of factors (1980) that are used in the program placement equation in a two-stage estimation procedure, as in Rosenzweig and Wolpin (1986) and Pitt, Rosenzweig, and Hassan (1990). Because the latent factors for the four age- and gender-specific school attendance groups are highly correlated, no precise estimation of the individual school attendance factors was possible. Statistical tests indicate that the set of four can be reduced to any two. Accordingly, results are reported using two schooling factors and those for fertility and child mortality.

Table 7 reports the weighted two-stage least squares estimates for each of the program variables of the effects of the kecamatan characteristics on kecamatan

Table 7. *Weighted Two-Stage Least Squares Estimates of the Determinants of Spatial Program Placement in 1980*

| *Variable* | *Grade school* | *Middle school* | *High school* | *Health clinic* | *Family planning clinic* |
|---|---|---|---|---|---|
| Latent fertility[a] | −0.120 | −0.082 | −0.060 | −0.048 | −0.120 |
| | (5.13) | (3.69) | (3.66) | (1.93) | (3.40) |
| Latent mortality[a] | −0.0061 | −0.197 | 0.099 | −0.390 | 0.950 |
| | (0.02) | (0.53) | (0.36) | (0.95) | (1.57) |
| Latent schooling, females ages 10–14[a] | −0.428 | 0.790 | 0.457 | −1.90 | 1.70 |
| | (0.51) | (1.02) | (0.79) | (2.18) | (1.36) |
| Latent schooling, males ages 10–14[a] | −0.618 | −0.620 | −0.200 | 0.930 | −2.30 |
| | (0.82) | (0.89) | (0.40) | (1.18) | (2.07) |
| Schooling of women ages 25–29 | 0.072 | 0.033 | 0.011 | 0.048 | 0.022 |
| | (7.20) | (3.61) | (1.63) | (4.59) | (1.49) |
| Schooling of mothers of children ages 10–14 | −0.100 | −0.083 | −0.033 | −0.018 | −0.093 |
| | (3.76) | (3.30) | (1.76) | (0.63) | (2.30) |
| Schooling of heads of households with children ages 10–14 | 0.037 | 0.067 | 0.046 | 0.0093 | 0.071 |
| | (1.84) | (3.61) | (3.33) | (0.44) | (2.34) |
| Proportion of households in urban areas | −0.063 | 0.410 | 0.430 | 0.220 | 0.170 |
| | (2.32) | (16.20) | (22.90) | (7.70) | (4.11) |
| Land owned by household ($\times 10^{-2}$) | −0.210 | 0.017 | 0.0031 | 0.010 | −0.160 |
| | (14.70) | (1.29) | (0.31) | (0.68) | (7.13) |
| Test statistics: Latent variable significance, $F(4,2537)$ | 29.8 | 10.2 | 7.65 | 11.1 | 12.2 |

*Note:* Specification also includes three altitude variables; indicators of history of earthquakes, drought, floods, and other natural disasters; and proximity of *kecamatan* to coast. Absolute values of asymptotic *t*-ratios are in parentheses.

a. Variable potentially measured with error. Instruments include 1985 latent variable measures.

program coverage as of 1980. For brevity, the five locational and physical-characteristics variables describing each kecamatan are omitted from the table. Tests of the joint significance of the effects of the four fixed latent factors on the placement of programs are reported at the bottom of the table. These statistics indicate that the placement of each of the five programs as of 1980 was significantly related to the unmeasured fixed factors relating to the six policy outcomes. Evidently, the distribution across kecamatans in the coverage of programs is not random with respect to the unmeasured factors determining outcomes and behaviors.

The estimates of the latent factor effects indicate that, in particular, kecamatans in which fertility is high, net of program and parental schooling effects, receive a lower level of program coverage with respect to all of the five programs or institutions. Most notably, kecamatans with a propensity to have higher fertility receive less family planning support, suggesting that such support is provided where it is most desired. It is less obvious why, net of the latent factors determining school attendance, the villages in high-fertility kecamatans are less likely to have schools. Because of the evident collinearity between the latent school attendance factors (which are jointly significant), it is not possible to discern the effects of those factors. The latent factors determining child mortality, however, do not appear to influence the coverage of any of the programs.

The results obtained by stratification of the population into women's schooling groups suggested that the targeting of grade schools to areas in which women (mothers) have lower levels of schooling would be more effective in raising average attendance rates for 10- to 14-year-olds than a random allocation would. The estimates in table 7 suggest that grade school coverage is lower, net of the influence of the latent school attendance factors, in kecamatans in which mothers of teens ages 10–14 have higher levels of schooling, consistent with efficiency criteria. Such areas, however, are also less likely to receive middle and high schools, for which no nonlinear effects were found. The negative relation between school placement and the schooling attainment of women may reflect equity concerns, although kecamatans characterized by (male) heads of households who have higher levels of schooling are more likely to receive each of the three levels of schools.

### *Is There Convergence in Program Coverage?*

As noted, in 1980 the six types of programs tended to be clustered, in that the programs were spatially positively inter-correlated. In this section, we report estimates of the dynamic version of the placement "rules," equation 4, which assess how the change in program coverage across kecamatans between 1980 and 1986 is related to the latent outcome effects, the measured time-invariant characteristics of subdistricts, the 1980 population characteristics, and the 1980 program coverage. Estimates of the effects of the 1980 program distribution on the growth in program coverage across kecamatans permit an assessment of whether between 1980 and 1986 program placement became more or less evenly

Table 8. *Weighted Two-Stage Least Squares Estimates of Growth in Program Coverage, without and with Correction for Measurement Error in Programs, 1980–86*

| | *Grade school* | | *Middle school* | | *High school* | | *Health clinic* | | *Family planning clinic* | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Variable* | *Without* | *With* | *Without* | *With* | *Without* | *With* | *Without* | *With* | *Without* | *With* |
| Grade schools (1980)[a] | −0.790 | −0.520 | −0.031 | 0.0160 | −0.032 | 0.0065 | −0.016 | 0.024 | 0.043 | 0.061 |
| | (44.8) | (13.2) | (0.84) | (1.91) | (1.06) | (0.10) | (0.35) | (0.22) | (0.74) | (0.69) |
| Middle schools (1980)[a] | 0.034 | 0.075 | −0.250 | 0.170 | 0.250 | 0.540 | 0.330 | 0.710 | 0.070 | −0.020 |
| | (2.60) | (1.54) | (9.17) | (1.62) | (11.4) | (6.91) | (9.86) | (5.35) | (1.65) | (0.17) |
| High schools (1980)[a] | 0.049 | 0.023 | 0.094 | 0.070 | −0.460 | −0.350 | 0.065 | 0.0054 | 0.031 | 0.120 |
| | (2.91) | (0.43) | (2.69) | (0.34) | (16.2) | (4.13) | (1.51) | (0.04) | (0.57) | (0.94) |
| Health clinics (1980)[a] | −0.036 | −0.034 | 0.077 | 0.014 | 0.052 | 0.020 | −0.590 | −0.300 | 0.027 | −0.097 |
| | (1.86) | (0.77) | (1.90) | (0.15) | (1.57) | (0.27) | (11.90) | (2.41) | (0.43) | (0.89) |
| Family planning clinics | −0.003 | −0.220 | 0.033 | −0.450 | 0.0095 | −0.370 | 0.089 | −0.700 | −0.860 | −0.670 |
| (1980)[a] | (0.22) | (2.26) | (1.32) | (2.70) | (0.47) | (2.83) | (2.89) | (3.19) | (21.7) | (3.37) |
| Latent fertility[b] | −0.013 | −0.0083 | 0.061 | 0.044 | 0.033 | −0.0060 | −0.078 | −0.092 | −0.009 | −0.013 |
| | (1.24) | (0.58) | (2.75) | (1.46) | (1.85) | (0.26) | (0.25) | (2.34) | (0.25) | (0.38) |
| Latent mortality[b] | −0.120 | 0.350 | −0.210 | 0.630 | −0.460 | 0.350 | −0.610 | 1.00 | 0.550 | 0.850 |
| | (0.68) | (1.47) | (0.57) | (1.26) | (1.53) | (0.89) | (1.30) | (1.53) | (0.94) | (1.45) |
| Latent schooling, females | −0.630 | −0.730 | −0.570 | −0.810 | −0.048 | −0.410 | −0.900 | −0.790 | 0.840 | −0.660 |
| ages 10–14[b] | (1.50) | (1.26) | (0.65) | (0.67) | (0.07) | (0.43) | (0.83) | (0.50) | (0.61) | (0.46) |
| Latent schooling, males | 0.250 | 0.410 | 0.600 | 0.520 | 0.035 | 0.150 | 0.680 | 0.170 | −0.820 | 0.620 |
| ages 10–14[b] | (0.71) | (0.83) | (0.80) | (0.50) | (0.06) | (0.19) | (0.74) | (0.12) | (0.70) | (0.51) |

*Note:* Specification also includes three altitude variables; indicators of history of earthquakes, drought, floods, and other natural disasters; and proximity of *kecamatan* to coast. Correction for measurement error uses the 1976/77 program coverage information to construct instruments for the 1980 program coverage. Absolute values of asymptotic *t*-ratios are in parentheses.

a. To correct for measurement error, instruments include program distribution variables in 1977.

b. In the specifications with and without correction for measurement error, this variable is treated as measured with error. Instruments include 1985 latent variable measures.

distributed, that is, whether there was convergence or divergence in program coverage across kecamatans. If the coverage of a particular program grew less in kecamatans in which the coverage was already relatively high, then convergence is indicated.

The regression of a change in a variable on its initial level, required to test for convergence or equalization, is problematic because if the variable is measured with error, then it is easy to show that the coefficient of the initial program-level effect will be biased negatively. Thus one may falsely accept the convergence hypothesis. To eliminate this problem, the program coverage information from the 1976–77 FASDES is used to construct instruments for the 1980 program coverage. As discussed in the appendix, we were able to match at the kecamatan-level data from this source, which provides similar program coverage information for the same variables, with the 1980 PODES data. If the measurement errors in the programs are independent across the two data sources, then the instrumented estimates are consistent.

Table 8 reports the weighted, two-stage, least squares estimates of the program coverage growth equation. Two sets of estimates are reported for each program. The first set was obtained without using the 1976–77 instruments to correct for measurement error in the 1980 program variables; the second set was obtained using the instruments and treating the 1980 programs as potentially error-ridden. On the basis of Hausman-Wu type tests, the hypothesis that the programs are measured without error is rejected in all cases except for the family planning growth equation.[10] The existence of the measurement-error problem is evident in the difference in the "own" effects of programs on their growth—for each program, there is an increase in the own effect when the instruments are employed. Indeed, for middle schools, the use of the instruments changes the sign of the effect of the 1980 coverage of middle schools on the 1980–86 growth rate of middle school coverage from negative (with an asymptotic $t$-ratio of more than 9) to positive (but not statistically significant). All of the other instrumented own-coefficient estimates, however, retain their negative sign and are statistically significant when the instruments are employed.[11] Thus, the results indicate that during 1980–86 program coverage across kecamatans was being equalized. Except for the middle schools, the coverage of a program grew less in areas that were better endowed with respect to that program.

10. The $F$-statistics (5,2098) for the grade school, middle school, high school, health clinic, and family planning clinic growth equations are, respectively: 95.4, 16.33, 8.82, 16.7, and 1.58.

11. The existence of measurement error in the programs, as indicated in the growth equations, implies that some of the estimates of program effects may be biased toward zero. It is not clear, however, what instruments are available to resolve this problem in estimating program effects. Models of dynamic decisionmaking imply that current decisions, such as the distribution of programs, depend on the values of current state variables (the 1980 program distribution in this case) and not on past-state values (unless they predict the future), so that past program distributions are valid instruments for 1980 programs in the post-1980 program growth equations. However, as discussed, the distribution of programs in the past may have a direct influence on cumulative human capital outcomes, given current programs.

The fixed-effects estimates of program effects indicated that health clinics had a positive impact on school attendance, net of the presence of schools, for females in both the 10–14 and 15–18 age groups. To the extent that these programs augment school attendance, fewer schools are needed to achieve the same attendance rates. Equalization would thus imply that grade schools and middle schools would grow less in kecamatans with a greater presence of health clinics. However, the estimates in table 8 indicate that the growth in the coverage of both school types was not responsive to the level of health clinic coverage in 1980. Moreover, the coverage of health clinics grew more in areas with a greater number of middle schools, as did the growth in coverage for high schools. One puzzling result in table 8 is that the coverage of each of the programs grew less in kecamatans with a higher coverage of family planning clinics in 1980; yet, there was no evidence of this program having any effects on the outcome variables.

## V. Conclusion

In this article, we have reported estimates of the effects of several important public programs associated with human resource investments (in schools, health clinics, and family planning clinics) on basic human capital indicators (school attendance, fertility, and child mortality). The estimates were based on a "new" data set constructed from a pool of kecamatan-level observations on human capital outcomes, socioeconomic variables, and program coverage based on the successive sets of cross-sectional household and administrative data describing Indonesia in 1976–86. This data set also enabled the investigation of the biases in conventional cross-sectional estimates of program effects arising from two sources: the lack of comprehensive information on programs and the nonrandom placement of governmental programs across areas. The data were also used to examine how the spatial allocation of programs in Indonesia in 1980 and the growth in program coverage by area were related to area-specific endowments in the 1980s and contributed to the efficiency of program effects and spatial and socioeconomic equity.

The empirical results, based on matched 1980 and 1985–86 information on more than 3,000 kecamatans, indicated that the presence of grade schools and, to a lesser extent, middle schools in villages has a significant effect on the school attendance rates of teenagers. The results also indicated that the presence of health clinics (*puskesmas*) in villages positively affects the schooling of females ages 10–18. Estimates based on the data stratified by the educational attainment of adult women also indicated that the effects of grade school proximity on the school attendance rates of teens ages 10–14 was significantly greater in households in which mothers had little or no schooling compared with households in which mothers had more than a grade school education. However, no other program effects appeared to differ across education classes of women. Moreover, based on the statistically preferred models, there was no evidence of any significant effects of the presence of family planning and health programs on either the survival rates of children or on cumulative fertility.

The estimates also suggested that the use of cross-sectional data, which does not take into account the possibly nonrandom spatial location of programs, results in substantial biases in the estimates of program effects because of the evident nonrandom spatial allocation of public programs. For example, cross-sectional estimates from the 1980 data resulted in an underestimate by 100 percent of the effect of being proximate to a grade school on the school attendance of both males and females ages 10–14, compared with estimates based on the pooled 1980 and 1985–86 data, which allowed for nonrandom program placement. The cross-sectional estimates also indicated clearly counterintuitive results, for example, that family planning clinics significantly raise fertility and reduce schooling investments. These results are not apparent when the nonrandomness of program placement is taken into account.

The estimates pertaining to the spatial and intertemporal allocation of programs in Indonesia indicated that the 1980 spatial distribution of each of the five programs examined here was significantly related to the unmeasured fixed factors relating to the six policy outcomes; the placement across kecamatans in the coverage of programs is not random with respect to the unmeasured factors determining outcomes and behaviors. Most notably, kecamatans with a propensity to have higher fertility received less family planning support, suggesting that such support is provided where it is most desired. The coverage of programs also tended to be lower in areas in which the educational levels of mothers was high, an allocation consistent with an efficiency criterion, given the finding that the effect of grade school proximity on school attendance is greater in households with less-educated mothers. However, this relation is also true for all of the programs studied for which there was no evidence of nonlinearities with respect to the schooling attainment of adult women. Finally, the examination of the change in the spatial allocation of programs between 1980 and 1986 indicated that the spatial program distribution became more equal; there was clear evidence of area-specific convergence in program coverage.

Although there was some evidence of significant program effects, particularly of school proximity on school attendance, it is apparent from exploiting the constructed longitudinal data that the quantitative estimates of these effects cannot account for a large part of the actual growth in human capital outcomes in Indonesia in the 1980s. In part this may be the result of measurement error in the program variables, on which there is some evidence, which would bias the program estimates toward zero. Some of the improvements in the human resource outcomes examined may reflect economic growth, which the data do not measure. Even with income information, however, the endogeneity of income must be considered as well as the possibility that human capital programs contribute to economic growth. Controlling for incomes could thus result in a misleading inference about the long-term consequences of public investments in human resource investments.

In future work we will explore the issue of the longer-term effects of these programs by extending the data set across time. Additional data will enable us to utilize the methods used here and to include the effects of lagged program distributions in the specification of program effects, to assess the role of area-specific income growth rates, and to investigate the effects of human resource programs on income growth.

## Appendix. Construction of the Data Set

To obtain a geographically consistent set of intertemporal observations, we matched the data set–specific geographic codes in two stages: the 1986 codes were matched to the 1980 codes, and the 1976 codes were matched to the 1980 codes. The 1980 and 1986 codes were matched using their respective master files. The province and district codes between two consecutive master files were matched, and then the subdistricts were matched by name. However, many names had changed or new subdistricts emerged because there were different abbreviations between periods or because some subdistricts split. The nonmatched subdistricts were then visually matched, but still the matching was not complete.

The subdistricts that were not matched based on names were brought to the attention of the Mapping Department at the BPS. From internal documents we tried to find the origin of the nonmatched subdistricts. However, the Mapping Department updated their maps in 1980 and 1986 only, not in 1983, and their documents listing code changes were not complete. For the remainder of the subdistricts that we could not match, we used various issues of the *Lembaran-Negara Republik Indonesia*, the annual, official government gazette in which decrees are published and which contains official documents recording villages changing subdistricts, new villages, new subdistricts, and boundary changes. This publication does not contain location codes, just names. To obtain the origin of the villages in the nonmatched subdistricts, we matched the village names from the gazette with village names from the master files. We changed the subdistrict codes according to the origin of most of the villages of the subdistricts in the master file. There were 103 location code changes between 1983 and 1986 and 217 location code changes at the subdistrict level from 1980 through 1983. Once we completed the master file changes, we converted the 1986 PODES into 1983 codes and then into 1980 codes.[12]

The 1976–77 FASDES contains just one code for the province and district combined, ranging from 1 to 287, whereas subsequently provinces and districts were identified with separate two-digit codes. To convert the FASDES geographic codes into 1980 codes, documentation on the three-digit location codes for the provinces and districts combined was used to update to the 1980 scheme of two-

12. We also tracked the location code changes from 1990 to 1986, as we had expected to be using data with 1990 location codes. From 1986 to 1990 we found 1,927 location code changes at the subdistrict level.

digit province codes and two-digit district codes. However, conversion of the FASDES codes to 1980 codes at the subdistrict level was made difficult by the fact that the FASDES subdistrict codes and names are not available. Thus, we had to match the village names from the FASDES master file along with their subdistrict codes with the village names and subdistrict codes on the 1980 master file. If five or more villages matched, we took the subdistrict codes from the 1980 master file. Village naming was sufficiently stable over time to permit the matching of all of the subdistrict codes between 1976 and 1980.

To convert the 1985 SUPAS into 1980 codes necessitated the use of the 1985 Sample List (*Daftar Sampel*), which contains the sample code numbers along with the province, district, subdistrict, and village codes. The raw data for SUPAS includes only the province and district codes, along with the sample code number. The three codes combined—for the province, the district, and the sample number—were used to obtain the subdistrict and village codes from the sample list. We converted the SUPAS into subdistrict codes using the sample list. These codes were based on the 1983 master file, so we then converted them from 1983 codes into 1980 codes.

Once the geographic codes of all of the data sets were made comparable, we aggregated the data at the common subdistrict level. With the PODES and FASDES, we calculated for each subdistrict the proportion of households whose village of residence contained each program, type of infrastructure, or environmental variable.[13]

Because the 1986 PODES was converted to 1980 codes, there were some duplicate location codes as villages and kecamatans split between 1980 and 1986. In 1986 there were 66,922 villages. Knowing which kecamatans and villages split between 1980 and 1986 allowed us to reaggregate 1986 administrative units back to their 1980 form. If areas were combined, we were, of course, unable to disaggregate program coverage into 1980 codes. There were 65,924 villages in 1986 with the 1980 codes. The FASDES did not contain any duplicate location codes after the conversion to 1980 codes. The total number of kecamatans in the 1980 PODES is 3,318. Of these, we were able to match all but 16 in the 1986 PODES.[14]

13. Although the PODES data are unique in the developing world for their detailed and comprehensive spatial information on program availability in 67,000 locales, they lack information on the "quality" dimension of public program provision. If quality is an important dimension of program effectiveness and if it is correlated with quantity measures of program availability, then estimation of a policy outcome equation will result in biased estimates of program effects. Even if quality and quantity are correlated, unbiased estimates of program effects can be obtained as long as variation in quality takes the form of a subdistrict-specific fixed effect or varies only with time. Again, this highlights the importance of data on program placement and outcomes at more than two points in time for the same village in eliminating bias.

14. Because the newly acquired province of East Timor was not included in the 1980 PODES, the 1986 PODES contains 68 more subdistricts.

## References


Gertler, Paul, and Jacques van der Gaag. 1990. *The Willingness to Pay for Medical Care: Evidence for Two Developing Countries.* Baltimore, Md.: Johns Hopkins University Press.

Huber, P. J. 1967. "The Behavior of Maximum Likelihood Estimates under Non-Standard Conditions." In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability,* 1.

Indonesia, Government of. Annual. *Lembaran-Negara Republik Indonesia.* Jakarta, Indonesia.

Pitt, Mark M., and Mark R. Rosenzweig. 1985. "Health and Nutrient Consumption across and within Farm Households." *Review of Economics and Statistics.* 67(2, May):212–23.

Pitt, Mark M., Mark R. Rosenzweig, and M. N. Hassan. 1990. "Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries." *American Economic Review* 80:1139–56.

Rosenzweig, Mark R., and Paul Schultz. 1982. "Child Mortality and Fertility in Colombia: Individual and Community Effects." In *Health Policy and Education,* 2. Amsterdam: Elsevier Scientific Publishing Co.

Rosenzweig, Mark R., and Kenneth I. Wolpin. 1982. "Governmental Interventions and Household Behavior in a Developing Country." *Journal of Development Economics* 10:209–25.

———. 1986. "Evaluating the Effects of Optimally Distributed Public Programs: Child Health and Family Planning Interventions." *American Economic Review* 76:470–82.

Strauss, John. 1990. "Households, Communities, and Preschool Children's Nutrition Outcomes: Evidence from Rural Côte d'Ivoire." *Economic Development and Cultural Change* 38(2):231–62.

White, Halbert. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48:817–38.