# Spatial Decentralization and Program Evaluation: Theory and an Example

Mark M. Pitt
Brown University

Nidhiya Menon
Brandeis University

Version: August 8, 2011

## Abstract

This paper proposes a novel instrumental variable method for program evaluation that only requires a single cross-section of data on the spatial intensity of programs and outcomes. The instruments are derived from a simple theoretical model of government decision-making in which governments are responsive to the attributes of places and their populations, rather than to the attributes of individuals, in making allocation decisions across space, and have a social welfare function that is spatially weakly separable, that is, that the budgeting process behaves as if it is multi-stage with respect to administrative districts and sub-districts. The spatial instrumental variables model is then estimated and tested by GMM with a single cross-section of Indonesian census data. The results offer support to the identification strategy proposed.

Keywords: Spatial Decentralization, Program Evaluation, Instrumental Variables
JEL Classification Number: C21, H44, O12, C50

---

Thanks to Andrew Foster and Moshe Buchinsky for comments on a much earlier version of this paper. Thanks also to participants at the 2007 North American Summer Meeting of the Econometric Society, and the 2010 Northeast Universities Development Consortium Conference. The usual disclaimer applies.
Contact information: Mark M. Pitt, Population Studies and Training Center, Brown University, Box 1836, Providence, RI 02912. Tel: 401.863.2445. Email: Mark_Pitt@brown.edu. Nidhiya Menon, Department of Economics & IBS, MS 021, Brandeis University, Waltham, MA 02454. Tel: 781.736.2230. Email: nmenon@brandeis.edu.

## 1. Introduction

Governments in developing countries earmark significant proportions of their budget towards establishing programs that seek to alter the behavior of target populations. By influencing fertility, health, and schooling outcomes, these programs are often the government's main tools for spreading economic well-being and for spurring economic growth. A fundamental problem in program evaluation is that the coverage of programs and the timing of program initiatives -- program placement -- are not likely to be random 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 estimates of program effects. Data on the spatial distribution of programs and population characteristics at more than one point in time can be used to identify program effects with relatively simple methods (fixed effects) when program placement depends on unmeasured time-persistent or permanent characteristics of locations but varies as a function of aggregate economy-wide trends or shocks. The longitudinal data required for fixed effects estimation are not always available in developing countries, or are too closely spaced so that program change is small relative to noise, and the assumption of the time invariance of the confounding unobservable component may not always hold. This paper proposes a novel instrumental variable method for program evaluation that only requires a single cross-section of data on the spatial intensity of programs and outcomes, does not require the strong assumptions of fixed effects methods, and can take into account time-varying area-specific heterogeneity (as evidenced in a cross-section of data) that cannot be dealt with by fixed effects methods that attempt to control for district-level or regional-level variation. The instruments are derived from a simple theoretical model of government decision-making that requires that the government's social welfare function is spatially weakly separable, that is, that the budgeting process behaves as if it is multi-stage with respect to administrative districts and sub-districts. The spatial instrumental variables model is then estimated and tested by GMM with a single cross-section of Indonesian census data.

Rosenzweig and Wolpin (1986) were among the first to examine the problem of the endogeneity of program placement using fixed effects methods. Using longitudinal data on nutritional status, they 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, while in fact they enhanced it once the endogeneity of the placement was "controlled." Pitt *et al.* (1993) estimate the effects of a number of important public programs - schools and health and family clinics - on basic human capital indicators - school enrollment, fertility and cumulative mortality rate of children - based on an Indonesian data set consisting of a pool of sub-district level observations on human capital outcomes, socioeconomic variables, and program coverage derived from successive sets of cross-sectional household and administrative data describing the period 1976-1986. These are the mostly the same programs and outcomes investigated in this paper. Their estimates also suggested that the use of cross-sectional data, which does not take into account the possibly non-random spatial location of programs, results in substantial biases in the estimates of program effects. For example, cross-sectional estimates resulted in an underestimate by 100 percent of the effect of being proximate to a grade school on the school enrollment of both males and females aged 10-14, compared to estimates based on the pooled data and fixed effects.

Duflo (2001) also examines the Indonesian case, making use of an unusual policy experiment, a massive school building program, to estimate the returns to school building on education and earnings. Using a cross-section of men born between 1940 and 1972, Duflo linked an adult's education and wages with district-level data on the number of new schools built between 1973-74 and 1978-79 in his region of birth. Conditional on cohort and region of birth fixed effects, interactions between dummy variables indicating age of the individual in 1974 and the intensity of the program in their region of birth are the instruments used in the wage equation, exploiting the variation in program intensity across regions and cohorts. The validity of these instruments requires that there are no omitted time-varying or regional effects correlated with the school building program, for example, the concurrent expansion of other programs as a consequence of the oil boom that funded the school program may invalidate these instruments.

The spatial instruments proposed in this paper are derived from models that optimize household and government behavior subject to resource and information constraints. The idea is that the nature of the budgetary process with spatially-defined administrative districts generates a set of exclusion restrictions that can be used as instruments for the observed allocation of public programs across space. Governments are, rather innocuously, assumed to be responsive to the attributes of places and their populations, rather than to the attributes of individuals, in making allocation decisions across space. If the attributes of a district rather than individual characteristics decide program placement, and the means (or higher moments) of outcomes for all "competing" districts enter into the government's social welfare function, then the district means of individual and district exogenous determinants of these outcomes in *other* districts may be used as instruments for the placement of programs in any particular district. The assumption of weak separability of a social welfare function having as argument the mean outcomes of every administrative unit (district) is sufficient to generate *spatially decentralized budgeting*, an allocation process that yields restrictions sufficient for identification. Decision-making of this type will arise as a consequence of the costliness of acquiring and processing information on the returns to and health status of every single uniquely identified household or spatial aggregation of households. The assumptions of this model, set out in detail below, are not specific to Indonesia or developing countries more generally. Although it simplifies consideration of the spatial allocation process to think of it as a process in which larger administrative districts allocate a budget to smaller districts who in turn allocate them to even smaller administrative units, **it not required that the government actually act in this manner,** only that the central government use multi-stage budgeting as in the usual case of separable utility in consumer demand theory applied to a single economic agent.

Section 2 of the paper sets out a theory of household and government decision-making that generates the required exclusion restrictions. Section 3 discusses issues of empirical implementation, and Section 4 describes the data and variable construction. In Section 5, models of program evaluation with spatial instruments are estimated by GMM using the 1980 Population Census of Indonesia merged with detailed information on the presence of government programs in all of the villages of Indonesia that were collected as part of the Population Census. Two sets of spatial instruments are constructed – one defined by spatial contiguity (shared boundaries) and one by membership in a larger administrative unit. Four outcomes (girls' and boys' school enrollment, recent fertility, and contraceptive use) and four programs (primary schools, secondary schools, family planning clinics, and public health clinics) are investigated. Hansen-Sargan tests of overidentification and tests of instrument orthogonality and instrument redundancy are reported. The instrumental variable results generate significantly different estimated program effects as compared to a model with exogenously placed programs. Section 6 summarizes the results.

## 2. Economic model

### Model for the household

To illustrate the sources of endogeneity confounding the evaluation of spatially sited programs and the assumptions underlying our identification strategy, we model both household and government behavior in the context of a multi-district nation. The household side is completely conventional. We abstract from issues of allocation within the household. Household *j* in sub-district *k* of district *l* has the following preference function:

(1) $U_{jkl}=U(H_{jkl},Z_{jkl})$

where $H_{jkl}$ denotes a health or human capital outcome, and $Z_{jkl}$ represents a composite purchased consumption good consumed by household *j*. Good $H_{jkl}$ is produced with production function

(2) $H_{jkl}=H(R_{kl}, F_{jkl}, E_{jkl}, \mu_{kl}, \eta_{jkl})$

where $R_{kl}$ denotes the level of public good inputs (e.g. public health services) provided by the government to sub-district *k* in district *l*, $F_{jkl}$ is an input chosen by the household, and $E_{jkl}$ represent exogenous environmental factors that are specific to the household (such as parental schooling and age) or specific to the district (climate, topography, geology). The term $\mu_{kl}$ is the sub-district specific unobservable which represents sub-district level heterogeneity, and $\eta_{jkl}$ is a non-systematic household specific error term representing deviations from sub-district average unobserved factors.

The household's budget constraint is given by:

(3) $I_{jkl}=P_{Zkl}Z_{jkl} + P_{Fkl}F_{jkl}$

where $P_{Zkl}$ and $P_{Fkl}$ are per unit prices of $Z_{jkl}$ and $F_{jkl}$ respectively. The reduced form demand equation (4) for $H_{jkl}$, conditional on $R_{kl}$ as well as prices, the environment, and the unobserved factors is obtained from the maximization of the utility function (1) subject to the production function (2) and the budget constraint (3):

(4) $H_{jkl}= h(P_{Zkl}, P_{Fkl}, E_{jkl}, R_{kl}, \mu_{kl}, \eta_{jkl})$

The linearized version of this reduced form demand equation is:

(5) $H_{jkl}= \beta_0 +\beta_1 P_{Zkl} + \beta_2 P_{Fkl} + \beta_3 E_{jkl} + \delta R_{kl} + \mu_{kl} + \eta_{jkl}$

The parameters δ, the effect of health public goods (*programs*) on the stock of human capital $H_{jkl}$, are the parameters of interest.

**Model for the social planner**

The country consists of *L* administrative districts each with *N* households, and *K* sub-districts per district each with $J = N/K$ households.[1] The pair of sub-district and district indices $(\kappa,\ell)$ uniquely identifies a sub-district. The central government has a social welfare function that includes many outcomes in addition to $H_{jk}$, the level of human capital enjoyed by its citizens. The "Ministry of Human Capital" (social planner) receives a lump-sum to allocate towards the production of human capital $H_{jkl}$. The most general form of the sub-utility function for human capital of the social planner contains the individual human capital outcomes of each household (person) as arguments:

[1] It is not required that there actually are administrative districts and sub-districts that have control over resource allocation in order for the process that allocates programs to places to be a spatially two-stage process as modeled here. It is sufficient that the central planner consider spatially aggregated program determinants and program outcomes in allocating programs that have spatially limited effects on outcomes. The spatial aggregation of the information set will arise as a consequence of the costliness of acquiring and processing information on the returns to and health status of every single uniquely identified household, as described below.

(6) $W=W(H_{111},H_{211},\ldots,H_{n11},H_{121},\ldots,H_{NKL})$

The cost per unit of operating the program *R* is *s*, so that the total cost of all programs in a sub-district is $sR_{kl}$. If V is the total amount allocated to the program, the government's budget constraint is

(7) $V = s\sum_l \sum_k R_{kl}$

Maximizing the social welfare function (6) subject to the budget constraint, normalizing s=1, and solving for $R_{kl}$, yields the reduced-form equations for program intensity

(8) $R_{kl} = R(P_{Z11},.., P_{ZKL}, \; P_{F11}, .., P_{FKL}, \; \mu_{11},.., \mu_{kl}, \; \{E_{jkl}\}, \{\eta_{jkl}\}, V)$

where $\{E_{jkl}\}$ and $\{\eta_{jkl}\}$ are shorthand for the list of length ($J$ x $K$ x $L$) household-specific values for observed ($E_{jkl}$) and unobserved ($\eta_{jkl}$) determinants of human capital for every household in the country.

There are two implications of equation (8) worthy of attention. First, the sub-district level error $\mu_{kl}$ is a determinant of both $H_{jkl}$, the health of individual *j* in sub-district *k* in district *l*, and of the program intensity in that sub-district, $R_{kl}$. Consequently, consistent estimation of the effect of program intensity on human capital must deal with this confounding variable. Second, the exogenous determinants of human capital in districts and sub-districts **other** than sub-district ($\kappa,\ell$), such as the prices $P_{Zkl}$ and $P_{Fkl}$, affect program intensity in sub-district ($\kappa,\ell$), $R_{\kappa\ell}$, but not human capital in ($\kappa,\ell$). These exogenous determinants might then seem on casual inspection to be available as identifying variables in the instrumental variable estimation of (5). Neither of these two implications requires that the program allocation process be characterized by decentralized or multi-stage budgeting, such as might arise from costly information and separable preferences, as discussed below. However, the values of the exogenous variables, such as $P_{Zkl}$ and $P_{Fkl}$, in sub-districts other than ($\kappa,\ell$) are **not** sufficient to identify the effect of programs in ($\kappa,\ell$) without imposing additional model structure, such as spatially decentralized budgeting. To see this, consider the linearized spatial program allocation equation, simplified to only include the two prices as observed arguments but otherwise fully parameterized:

(9) $R_{kl} = \pi^{kl}_{Z11} P_{Z11} + \pi^{kl}_{Z21} P_{Z21} + \ldots + \pi^{kl}_{ZKL} P_{ZKL} + \pi^{kl}_{F11} P_{F11} + \pi^{kl}_{F21} P_{F21} + \ldots + \pi^{kl}_{FKL} P_{FKL} + \upsilon_{kl}$

The number of sub-districts (*KL*) is, in general, far less than the number of parameters (2**KL***KL*) and thus this equation is cannot be estimated without restrictions on the parameters π.

What kinds of restrictions seem sensible but also permit the instrumental variable estimation of the effect of programs on $H_{jkl}$? The number of free parameters can be greatly reduced by assuming all spatial own effects are the same for all sub-districts and all spatial cross-effects are the same for all sub-districts,

$$(10) \qquad R_{\kappa\ell} = \pi_Z^c P_{Z\kappa\ell}^c + \pi_Z^o P_{Z\kappa\ell} + \upsilon_{\kappa\ell}$$

where only the terms involving $P_Z$ are written out to reduce clutter, and where the cross price term $P_{Z\kappa\ell}^c = (\sum_{l=1}^{L}\sum_{k=1}^{K} P_{Zkl})$ - $P_{Z\kappa\ell} = \bar{P}_Z$ - $P_{Z\kappa\ell}$, and the superscripts attached to the parameters π refer to own and cross-effects. The restrictions on equation (9) are that $\pi_{Zk'l'}^{kl} = \pi_{Zk''l''}^{kl}$ and $\pi_{Zkl}^{k'l'} = \pi_{Zkl}^{k''l''}$ for all (k',l') and (k'',l'') not equal to (κ,ℓ). These restrictions imply spatial symmetry on the part of the social planner – the marginal effect of a change in a price on program intensity is the same in any sub-district in which the change occurs, and the marginal change in a price in any sub-district has the same effect on other sub-districts irrespective of the pairs of sub-districts. A change in the determinants of the placement of public programs induces the same response by the social planner irrespective of sub-district. This is really no more than saying that all the observations can be pooled into a single regression function. It allows for the social planner to be lobbied, bribed, or otherwise induced to locate programs in a certain sub-district, and for the social planner to respond to that lobbying, requiring only that he respond the same manner irrespective of the identity of the sub-district which is the source of that lobbying.

The cross price term $P_{Z\kappa\ell}^c$ does not appear in the determinants of human capital equation (5) and thus may appear to be appropriate exclusion restrictions in identifying the effect of programs on human capital. Moreover, the first-stage equation is clearly estimable as the number of free parameters is, in this case, only two. However, $P_{Z\kappa\ell}^c$ is perfectly negatively correlated with the included variable $P_{Z\kappa\ell}$, so that there are, in fact, no identifying instruments in this specification without adding further structure.

**Spatially decentralized budgeting as a source of parameter identification**

The assumption of a spatially decentralized budgeting process provides the additional structure needed for identification. Spatially decentralized budgeting (i) keeps the number of parameters in the first-stage program allocation equation below the number of sub-districts, (ii) does not require that the investigator be informed *a priori* about the spatial preferences of the central planner, and (iii) does not yield perfectly correlated own- and cross-price variables.

The intent is to put credible restrictions on the very general social welfare function (equation 6) that are sufficient to identify the effects of the public programs R on human capital outcomes. First, it is assumed that the social welfare function can be written with sub-district level outcomes, such as the sub-district means or higher moments, in place of individual specific outcomes as arguments. This will arise as a consequence of the costliness of acquiring and processing information on the health status of every single uniquely identified household or small aggregations of households.[2] Decision makers may only have information from sample surveys on the human capital stock of its citizens and may only be aware of the first few moments of the sampled distribution in sub-districts, and have only these statistics in allocating resources.

The assumption of weak separability of the spatially defined good $H_{jkl}$ by district and sub-district aggregations is sufficient to generate *spatially decentralized budgeting*, an allocation process that yields restrictions sufficient for identification. The weak separability restriction says that the marginal rates of substitution between the human capital of any two sub-districts are independent of the human capital of all other sub-districts, hence aggregator functions exist for sub-districts. As in the case of consumer demand theory, the spatial allocation process with weak separability can be conceptualized as a two-stage budgeting process in which budget allocations are made to districts and then, conditional only on the budget available to the district to which a sub-district belongs, allocations are made to each sub-district. This process looks like (spatially) political decentralization, in which decision-making is devolved to local levels of government from larger units of government. Local units of government are likely to be better informed about the distribution of human capital outcomes and the health environment, and so devolving decision-making to them may be efficient when conveying this information is expensive or fraught with strategic misrepresentation. However this may be, whether or not decision-making is actually devolved to lower-levels of government does not matter for the estimation strategy, only that the social welfare function is weakly separable.

Weak separability for subsets of human capital outcomes defined by common sub-districts of residence yields the social welfare function

$$(11) \qquad W=W(w_1[w_{11}(H_{11}),w_{21}(H_{21}),\ldots,w_{K1}(H_{K1})],\ldots, w_L[w_{1L}(H_{1L}),w_{2L}(H_{2L}),\ldots,w_{KL}(H_{KL})]$$

where $H_{kl}=\{H_{1kl},\ldots,H_{nkl}\}$, $w_l$ is a sub-utility function for district *l* defined over the second-level sub-utilities $w_{kl}$ of all of the sub-districts in *l*.[3]

[2] In addition, the social planner's allocation decisions only affect the sub-district "prices" of human capital inputs – the availability of schools and clinics-- and not person-specific prices.

[3] The social welfare function used above to motivate the exclusion restrictions arising from multi-stage budgeting posited an apparently benevolent social planner who derives welfare from the well-being of citizens. This assumption is not necessary to generate the exclusion restrictions. The social planner can be motivated by electoral or other political considerations. All that is required is that the informational complexity of weighting the political characteristics of all individuals requires some level of spatial aggregation consistent with multi-stage budgeting.

Consider multi-stage budgeting (separable preferences) in which the central planner allocates resources for the program to districts, and the *district social planner* allocates resources to sub-districts, as in equation (11). The allocation of program intensity to sub-district ($\kappa,\ell$) depends on two different own and cross-price influences by which sub-districts "compete" for resources from the central planner and the district planner. First, at the topmost level of budgeting, the attributes of district $\ell$ are compared to all other districts $l\neq\ell$. Second, at the second-level of budgeting, the attributes of sub-district $\kappa$ in district $\ell$ are compared to that of all other sub-districts in district $\ell$. The linearized allocation equation with two-stage budgeting is of the form

$$(12) \qquad R_{\kappa\ell} = \pi^{c}_{\mathrm{ZL}} P^{c}_{ZL(\kappa\ell)} + \pi^{c}_{\mathrm{ZK}} P^{c}_{ZK(\kappa\ell)} + \pi^{o}_{\mathrm{ZL}} P_{Z\ell(\kappa\ell)} + \pi^{o}_{\mathrm{ZK}} P_{Z\kappa\ell} + \upsilon_{\kappa\ell}$$

where $P^{c}_{ZL(\kappa\ell)} = \sum_{l\neq\ell}\sum_{k=1}^{K} P_{\mathrm{Zkl}}$, that is, prices $P_{\mathrm{Zkl}}$ summed over all sub-districts outside of district $\ell$, $P^{c}_{ZK(\kappa\ell)} = \sum_{k\neq\kappa} P_{Zk\ell}$, that is, prices $P_{\mathrm{Zkl}}$ summed over all sub-districts inside of district $\ell$ but excluding sub-district $\kappa$ in that district, and $P_{Z\ell(\kappa\ell)} = \sum_{k=1}^{K} P_{Zk\ell}$. These aggregations of prices can, of course, be rescaled to means. The variable, $P^{c}_{ZL(\kappa\ell)}$, the (mean) price of all other districts in the country, is the cross-effect at the first-level of budgeting, and the variable $P^{c}_{ZK(\kappa\ell)}$ is the cross-effect of all other sub-districts in district $\ell$ in the second-level of budgeting. The variable $P_{Z\ell(\kappa\ell)} = \sum_{k=1}^{K} P_{Zk\ell}$, the (mean) prices of all sub-districts within district $\ell$, and $P_{Z\kappa\ell}$, are the own-effects in the first- and second stage of budgeting, respectively. As before, note that collinearity arising from

$$(13) \qquad P^{c}_{ZL(\kappa\ell)} = \bar{P}_Z - P_{Z\kappa\ell}$$

and

$$(14) \qquad P^{c}_{ZK(\kappa\ell)} = P_{Z\ell(\kappa\ell)} - P_{Z\kappa\ell}$$

implies that the set of parameters $\pi$ are not all separately identified. Rewriting (12) in terms of the identified parameters, and omitting the constant term $\pi^{c}_{\mathrm{ZL}}\bar{P}_Z$, the estimated first-stage equation is:

$$(12') \qquad R_{\kappa\ell} = (\pi^{c}_{\mathrm{ZK}} + \pi^{o}_{\mathrm{ZL}})\, P_{Z\ell(\kappa\ell)} + (\pi^{o}_{\mathrm{ZK}} - \pi^{c}_{\mathrm{ZL}} - \pi^{c}_{\mathrm{ZK}})\, P_{Z\kappa\ell} + \upsilon_{\kappa\ell}$$

The second-stage budgeting assumption provides the source of identifying variation in equation (12'), the variable $P_{Z\ell(\kappa\ell)}$, the (mean) price of good Z in district $\ell$ that is not found in equation (10). The validity of this instrument requires that the exogenous attributes of places in the larger administrative unit to which a sub-district belongs affect the allocation of programs to that sub-district (two-stage budgeting), and that the exogenous attributes of other sub-districts not directly affect human capital outcomes in that sub-district.

**Identification using own sub-district attributes only: the Hausman-Taylor method**

If individual-level data on program outcomes are available in the data, then the instrumental variable estimation method of Hausman and Taylor (1981) can be applied without having to make multi-stage budgeting (weak separability) assumptions on the allocation process of the social planner. So far we have focused only on sub-district specific prices of human capital inputs and consumption goods ($P_{Zkl}$ and $P_{Fkl}$, respectively) to motivate the exclusion restrictions that arise from multi-stage budgeting of programs. Our health production function also includes household-specific exogenous attributes $E_{jkl}$ that appear in the program allocation reduced form equation (5). Unlike sub-district specific exogenous variables like prices, sub-district means of household-specific exogenous variables (such as parental schooling and age) are valid instruments for program placement without assuming decentralized budgeting, but require other assumptions. To simplify the exposition, we abstract from prices and again consider a linear program allocation equation in which the social planner allocates program resources in response to the sub-district means $E_{\kappa\ell} = \sum_{j=1}^{J} E_{j\kappa\ell}$ of the household-level exogenous variable affecting human capital,

$$(15) \qquad R_{\kappa\ell} = \pi_E^c E_{\kappa\ell}^c \ + \ \pi_E^o E_{\kappa\ell} \ + \ \upsilon_{\kappa\ell}$$

where $E_{\kappa\ell}^c = \ (\sum_{l=1}^{L}\sum_{k=1}^{K} E_{kl})$ - $E_{\kappa\ell} = \bar{E}$ - $E_{\kappa\ell}$, that is, the characteristics of all other sub-districts except ($\kappa$,$\ell$). As in the case of prices, $E_{\kappa\ell}^c$ is perfectly collinear with $E_{\kappa\ell}$, but in contrast to the case of prices, the fact that only the sum of the own ($\pi_E^o$) and cross-effect ($\pi_E^c$ ) is identifiable is not an impediment to identification of program effects if household-level data are available. This is a consequence of the assumption that only individual-level environmental variables ($E_{jkl}$) (such as parental age and schooling) are determinants of individual-level human capital conditional on programs in (5).

In the case of Hausman-Taylor (*H-T*) identification, the assumption required is that of the independence of individual-level human capital levels and the attributes of one's neighbors in the sub-district. For example, it requires that the age and schooling of a child's parents influence that child's human capital but not the human capital of other children in a sub-district conditional on public programs. In the case of investment in schooling, if labor markets are spatially separated and spatial mobility is costly, then the characteristics of parents affect wages and the return to those investments in the sub-district. There are other scenarios under which this *spatial independence* assumption is not valid. If there are neighborhood externalities to health, such as in the case of de-worming (Miguel and Kremer, 2004) then spatial independence may not hold. If there are peer group effects or other types of non-independent preferences within spatially defined areas, then spatial independence will also not hold. We test and reject the validity of the Hausman-Taylor estimator as applied to our data below.

**Sources of identification summarized**

To summarize, a model of the placement of spatially sited public programs affecting human capital in which the social planner's welfare function is defined over the human capital outcomes of all spatially defined districts, yields identifying instrumental variables (exclusion restrictions) for estimating the effects of public program on human capital outcomes of the following types:

*A.* *Multi-stage budgeting*. Means of variables that only vary across spatially defined districts, such as prices and the within-district means of household-level variables, are available as instruments if decision-making is spatially decentralized and districts "compete" with each other for resources from the central social planner and (possibly) district-level social planners in subsequent stages of the allocation process.

B. *Hausman-Taylor identification*. Sub-district means of household-level determinants of human capital (such as parental schooling and age) can serve as instruments under the assumption that the mean of these characteristics are uncorrelated with the unobserved sub-district-level effect.

## 3. Issues in empirical implementation

The spatially decentralized (multistage) budgeting model that we propose means that, in the terminology of Conley and Ligon (2002), *economic distance* matters. What characterizes the economic distance relevant to program placement is not obvious. The theoretical example discretizes economic distance by characterizing it as whether administrative sub-districts ($\kappa,\ell$) and ($\kappa',\ell'$) are both contained within the same district ($\ell = \ell'$). This notion is more appropriate to formally decentralized program allocation in which decision-making is devolved to sequentially lower units of government than to a single social planner. The use of membership of a place (sub-district) in a district as the measure of economic distance has the implication that two places that are spatially contiguous but in different districts (they straddle the district border) are as economically distant as places at the opposite ends of a country. Other general metrics include physical distance (distance between centroids of sub-districts) or even travel time distance. In the empirical example using data from Indonesia presented below, we lack data on physical distance. Instead, we construct two measures of economic distance, one based on spatial proximity and one based on shared district status. In the *neighbors* measure, sub-districts that are contiguous are considered "spatially proximate" irrespective of their district. In the second measure, labeled *non-neighbors*, membership in a common district is what matters. In the case of neighboring districts (contiguity), some but not all of the other sub-districts in district $\ell$ are likely to be contiguous with ($\kappa,\ell$), and sub-districts outside of district $\ell$ may be contiguous with ($\kappa,\ell$). The idea is that competition for public programs with neighboring sub-districts with which a sub-district shares program effects across a common border may differ from competition with non-neighboring sub-districts with which, by dint of membership in a larger political unit, it shares an allocation for a program to be spatially distributed by the district (*kabupaten*) social planner. To avoid overlap and for other reasons discussed below, the *non-neighbors* designation refers to sub-districts in the same district as ($\kappa$,$\ell$) that are not neighbors with it. Thus the union of *neighbors* and *non-neighbors* is all other sub-districts in district $\kappa$ except ($\kappa$,$\ell$), plus any sub-districts contiguous to ($\kappa$,$\ell$) that are not in district $\kappa$.

An additional consideration arises if the benefits of spatially sited public programs can spill over the boundaries of spatially defined sub-districts. In principle, programs can serve clients from a catchment area larger than a single sub-district. Given an existing spatial distribution of hospitals or secondary schools, for example, the social planner may be less likely to invest in a new facility in a particular district if that district is already served by similar facilities in nearby districts. If population was distributed uniformly across a homogenous plane and administrative boundaries did not affect access to program services, an efficient planner would locate new facilities depending only on distance to existing facilities. If a facility like a hospital or secondary school can serve a group of spatially proximate sub-districts, its location depends on the attributes of those spatially proximate districts differently than it does on less proximate districts. Consequently, the attributes of neighboring districts may have an especially large effect on whether a program is sited in a particular district. More importantly, it renders the exogenous characteristics of *neighbors* invalid as exclusion restrictions in the estimation of the human capital equation (5) conditional on own-sub-district programs. We use the exogenous characteristics of *non-neighbors* as an additional instrument set to test for the orthogonality of *neighbors* with respect to the unobserved components of equation (5).

A common approach to estimating a policy response function such as (5) is to condition on a single policy, for example increasing the spatial coverage of a single program from the set of programs $R_{kl}$. This single-program policy response function provides the effect of increasing the single program's "intensity" (say schools) on the outcome of interest (school enrollment, as in Duflo (2001)). However, it is clear that, in general, the spatially defined instruments (including those of the Hausman-Taylor type) may, in general, affect the social planner's allocation of the full set of programs $R_{kl}$, and that the spatial allocation of individual programs may be correlated.[4] If these other programs also affect the outcome of interest, for example fertility control and public health programs may affect schooling of children in a general model of child quantity and quality, the orthogonality conditions underlying the instruments will not hold. Consequently, this approach may only be valid when estimating a policy function (5) that condition on the complete

[4] Pair-wise correlation coefficients for the four government programs in analysis are consistently large, positive and significant at the 5% level., suggesting that there are likely to be complementarities in the spatially placement of program and in their effect on individual outcomes.

set of programs. This is testable with the data as long as sufficiently large set of instruments are available, and we report these tests below.

Finally, the spatial decision-making process that we posit implies that unobserved attributes that affect human capital outcomes and spatially sited public programs are likely to be correlated across space, and thus the standard errors that we construct for our instrumental variables models account for this non-independence by clustering at the district level.

## 4. Data and variable construction

We use data from two sources in this study: the 1980 *Potensi Desa* (Village Potential) survey of Indonesia (*PODES*) and the 1980 *Sensus Penduduk* (Population Census) of Indonesia. The 1980 *PODES* data provides information at the village level on the government programs that are studied: health clinics (*PUSKESMAS*[5]), family planning clinics, and grade and secondary schools. Information on area specific geographical characteristics such as the occurrences of natural shocks (droughts, floods, earthquakes, and other shocks) in the last five years, and other information such as distance from the coast and proportion of households in urban areas, is also reported. It is important to note that these natural shocks are typically not catastrophic events. For example, Indonesia experiences many earthquakes annually that result in little more than collapsed bridges and ruptured irrigation dikes. Records on earthquakes from 1917 to 2010 with an intensity of 6.0 (medium intensity) or higher show that there were 10 such earthquakes in 2005 and 2009, with a median of 0 fatalities in either year.[6] Approximately 62,000 villages (*desa*),

[5] PUSat KESehatan MASyarakat – literally, Peoples Health Centers.

[6] This information is available from the U.S. Geological Survey at http://earthquake.usgs.gov/earthquakes/world/historical_country.php. In terms of significant earthquakes, where earthquakes are classified as being significant if they caused moderate damage (about $1 million or more), 10 or more fatalities, had an intensity of 7.5 or higher, a modified Mercalli Intensity of X or higher, or the earthquake generated a tsunami, data for Indonesia from the National Geophysical Data Center of the National Oceanic and Atmospheric Administration shows that many of these instances occurred in fairly narrow geographic areas on land, and in retrospective data from 1629 to 2010, the median number of recorded total deaths is only 30 (many occur out at sea, and many do not create tsunamis large enough to cause fatalities). These data are available at National Geophysical Data Center/ World Data Center (NGDC/WDC) Significant Earthquake Database, Boulder, CO, USA. (Available at http://www.ngdc.noaa.gov/nndc/struts/form?t=101650&s=1&d=1). The earthquake and tsunami of December 26, 2004 is an outlier, since it is the only instance when deaths from a natural event in Indonesia exceeded the approximate 10,000 mark, a level that was previously set for the country in 1815.

almost all of the villages of Indonesia, are covered by the 1980 PODES data, which was carried out in conjunction with the 1980 Population Census.[7]

The 1980 long-form Census data sample provide detailed individual level information on the dependent variable outcomes: current school enrollment for girls and boys, and women's birth histories and contraceptive use. Data are also provided on other individual and household characteristics such as age and schooling attainment of household heads and spouses, as well as area of land owned by the household, indicators for whether the household owns its own home, religion of the household, and language of the household head. The sub-district level data was constructed by merging the information on programs and other geo-physical variables from the 1980 *PODES* data with the 1980 Indonesian census data aggregated to the *kecamatan* (sub-district) level. For each outcome, we have a random sub-sample of between 80,000-95,000 individuals who are "at risk" for the outcome. In particular, our sample consists of 82,891 girls ages 10-18, 82,889 boys ages 10-18, 87,655 women ages 21-30 with fertility histories, and 95,372 women ages 21-30 with recent contraception data.

Table 1 presents the means and standard deviations for each of the four outcomes analyzed in this study. In particular, the outcomes we study are current school enrollment rates of girls and boys ages 10-18 years, whether last child's year of birth lies in the previous two years (between 1978-1980) for women between the ages of 21 and 30, and whether any contraceptives are currently being used by women ages 21-30. Table 1 shows that current school enrollment rates range from 59% for girls 10-18 years of age to 66% for boys 10-18 years of age. Fertility, as measured by the incidence of births in the previous two years for women ages 21-30 years, is relatively high at 69%. In keeping with this, contraceptive prevalence among women in this age group is relatively low at 28%. The samples that are used to study the four outcomes we consider include data from approximately 3000 sub-districts in Indonesia. Table 1 also presents summary statistics for government programs. Approximately 77% of households reside in villages (*desa*) in which there is a grade school, and about 40% live in village with a junior or secondary school. Coverage of *PUSKESMAS* health clinics is about 25% in our samples, and the coverage of family planning clinics is approximately 49% in the data.

The exogenous variables are classified as follows:

1. <u>Variables describing sub-district (κ,ℓ) that vary across but not within sub-districts</u>. These describe the physical and economic environment of a sub-district (*kecamatan*) and correspond to the prices of the theory. These include measures of the recent (five-year)

[7] It is not possible to link village data from the *PODES* dataset to household-level data from the Population Census at the *desa* level. Although *desa* translates to village in common usage, every part of Indonesia is part of a *desa* in these data; *desa* are units of governance with well defined boundaries so that the largest cities contain hundreds of them and even smaller towns will have more than one.

occurrence of four types of natural shock (droughts, floods, earthquakes, and other natural shocks), whether the sub-district is located on an ocean coast, and the proportion of households in villages of the sub-districts that are urban. These sub-district variables are determinants of individual human capital outcomes for residents of that sub-district. The exclusion restrictions based upon multi-stage budgeting are that the sub-district means of these variables in competing sub-districts influence program placement in a particular sub-district but not human capital outcomes in that sub-district conditional on program placement.

2. Variables that vary within sub-districts. These are the household and individual variables that are determinants of the individual human capital outcomes, including parental age and schooling, area of land owned, religion, and the language of the household head, plus interactions of some of these variables with the proportion of the sub-district population that is urban. These variables are included in all regressions although their estimated parameters are not reported in the tables.

3. Hausman-Taylor identifying instruments. These are the sub-district means of the individual exogenous variables of #2 above. The underlying individual-level variables are presumed to affect individual human capital outcomes but the sub-district means of these household-level variables are assumed to influence only program placement in a sub-district (not individual human capital outcomes).

4. *Neighbor* instruments. These are the means of the variables of #1 and #3 taken over the sub-districts that are spatially contiguous to sub-district ($\kappa$,$\ell$). For each sub-district, we determine neighbors that share a geographical boundary using detailed province-level maps of Indonesia. There are at most 14 neighboring sub-districts in the data, although most sub-districts in the data have many fewer neighbors.[8]

5. *Non-neighbor* instruments. These are the means of the variables of #1 and #3 taken over the sub-districts that are in the same district (district $\ell$) as sub-district ($\kappa$,$\ell$) but are not spatially contiguous to sub-district ($\kappa$,$\ell$). Districts are known as *kabupaten* in Indonesia and there are approximately 300 *kabupaten* in the data. Instrument sets #4 and #5 arise from the weak separability of the social welfare function.

Table 2 reports descriptive statistics for the exogenous individual and household characteristics, as well as the interactions of these characteristics with proportion of households in the sub-district located in an urban zone as defined by the Population Census, plus the standard deviations for all variables that make up the *neighbor* and *non-neighbor* instrument set. In principle, higher moments of the distribution of sub-district characteristics, or other measures of

---

[8] For small island sub-districts, we define neighbors as the spatially closest sub-districts on the nearest islands.

central tendency such as medians, could be use in the spatial aggregations. Means are what are reported in the statistical reports of the Indonesian Central Bureau of Statistics and of government ministries, and thus are particularly relevant. There seems to be plenty of variation in spatial program placement attributable to the instruments derived from sub-district means, and adding higher moments may only add weak instruments. Table 2 also presents descriptive statistics for environmental variables that are measured at the sub-district (*kecamatan*) level. From this table, about 13% of households are in urban areas, and the proportion of households that have experienced natural shocks such as droughts, floods, or other events, ranges between 9% - 26% in these data. All of the own sub-district variables, including means of individual-level characteristics, listed in Table 2 are included in the second stage of all of the spatial instrumental variable models reported.

## 5. Results

We start by estimating Hausman-Taylor regression models for all four outcomes using the individual-level data. In Table 3, estimates presented in columns (1), (3), (5) and (7) condition only on the program or programs most associated with the particular outcome; school availability for school enrollment, and family planning clinics for recent fertility and contraceptive use. Estimates presented in columns (2), (4), (6), and (8) condition on all four programs. The heteroskedasticity consistent and spatially clustered Sargan-Hansen test statistics imply that the tests of overidentifying restrictions reject their null hypothesis in all of the models with a subset of the programs, and only fail to reject (at the 0.05 level) in the case of recent fertility with the full set of programs. As expected, adding programs to the specifications increases the p-value of this test statistic in every case. Nonetheless, these test statistics give us no confidence that the Hausman-Taylor instrument set is appropriate. [9] The results suggest that additional schools increase school enrollment, additional family planning clinics reduce fertility and increase the use of contraception, public health clinics (*PUSKESMAS*) reduce school enrollment, increase fertility and decrease contraceptive use, and primary schools reduce fertility and increase contraceptive use.

The patterns in Table 3 are qualitatively very similar to that found in Table 4 that estimates models that treat program placement as exogenous (OLS). As is well understood, OLS models

[9] The failure of the Hausman-Taylor instruments may reflect a Tiebout (1956) type of world in which households sort themselves spatially among communities offering different mixes of public services and other attributes. Household that highly value quality schools may themselves be highly educated and have births at later ages, and sort themselves into communities with households having similar preferences and personal attributes. In this example, the age and schooling of parents in neighboring and non-neighboring districts are not likely to be as strongly affected by Tiebout-like spatial sorting of households than they would be within districts. Robustness checks along this line are reported below.

that measure linear associations between program placement and outcomes, are not attentive to unobserved, location-specific attributes, and are thus unlikely to provide correct estimates of program impacts (Pitt *et al.* 1993). Facing resource constraints, governments in developing countries are likely to allocate programs to regions where the anticipated location-specific returns are high (*altruism theories*). This, in turn, may be influenced by the demographic and physical endowments of locations. Alternatively, *pressure-group theories* suggest that governments allocate a large share of program resources to populations with high demand for outcomes. The placement of new programs is also likely to be responsive to the presence of other identical programs that have already been placed in a region, since program-specific payoffs are likely to fall as similar programs are increasingly placed within the same region. Furthermore, returns to a program may also be influenced by the characteristics of the target households. For example, households with higher levels of schooling may derive greater benefits from public health programs.

For these reasons, OLS models which do not control for the non-random nature of program allocation across space result in biased estimates. In fact, if, as is likely to be the case, maximizing returns on a per-program basis underlies the geographic distribution of public funds, OLS will underestimate impacts. This is evident from the discussion and results in Pitt *et al.* (1993), and can be clearly seen on comparing the size of the OLS coefficients in Table 4 with the magnitude of the preferred estimates in Tables 5 through 8 which are discussed next.

Tables 5, 6, 7, and 8 present the GMM estimates of models estimated with sub-district level data using the instrument sets derived from the multi-stage budgeting model of program allocation. [10] Column (1) of Table 5 uses the *neighbor* sub-district instruments and conditions on only school availability in the determination of girl's school enrollment. Column (2) adds family planning clinics and *PUSKESMAS* to the specification. Hansen's J-test fails to reject the null hypothesis in either specification at the 0.10 level (p=.12 and p=.21, respectively), providing some confidence in the validity of this instrument set. The failure to reject the model that conditions on only school availability means that our estimation strategy permits the estimation of the usual sort of policy function – the effect of schools on schooling – without having to include other public programs in order not to reject the test of overidentification. The estimated effect of secondary schools on girl's school enrollment is about one-third larger when the two non-school programs are added to the specification. As this is a linear probability model, a 1 percent point increase in secondary school availability is estimated to increase school enrollment rate of girls ages 10-18 by 0.62 percentage points in the model with only schools, an effect that is almost 8

[10] The parameters on programs and their standard errors, as well as the Hansen-Sargan J-tests, are invariant to whether individual-level data or sub-district level data are used because the programs vary at the sub-district level. In order to estimate program effects there is no reason to use the individual-level data unless using Hausman-Taylor methods, in which case they are required.

times larger than the exogenous program model of Table 4.[11] Columns (5) and (6) replicate the specifications of columns (1) and (2) except that they present estimates of a grouped probit model by GMM. The Hansen J-test statistics are slightly smaller in the grouped probit model (p=.27 and p=.45 for the schools alone and "all programs" models, respectively). The marginal effect of secondary schools on girl's school enrollment is somewhat larger (0.87 versus 0.62) in the grouped probit model of column (5) as compared to the (grouped) linear probability model of column (1).[12]

In columns (3) and (4) of Table 5, the *neighbor* instrument set is augmented with the *non-neighbor* instrument set. A *C* or *GMM distance test* (Baum and Schaffer, 2007) is used to test the validity of the *neighbor* orthogonality conditions that underlie the GMM estimator of columns (1) and (2). Denote J as the value of the GMM objective function for the efficient GMM estimator that uses the neighbor plus the non-neighbor augmented instrumental variable set orthogonality conditions, and $J_N$ as the value of the efficient GMM estimator that uses only the *non-neighbor* orthogonality conditions, then under the null that the *neighbor* orthogonality conditions are actually satisfied, the test statistic $(J - J_N) \sim \chi^2$ with degrees of freedom equal to the number of variables in the *neighbor* instrument set. The null hypothesis that the orthogonality conditions associated with the neighbor instrument set is satisfied is not rejected in the model with only school programs in column (3) (p= 0.58) and in the model with all four programs in column (4) (p=0.42). In addition, we estimate an LM version of the Kleibergen–Paap (2006) rk test (Kleibergen and Schaffer, 2007)of the redundancy of the *non-neighbor* instrument set by testing the rank of the matrix E(X'Z). The null hypothesis that the non-neighbor instrument set is redundant when added to the neighbor instrument set is clearly rejected in both column (3) (p=.000) and column (4) (p=.000). The J-test of overidentification rejects the null hypothesis with the augmented instrument set whereas it does not reject with the *neighbor* instrument set alone, suggesting the appropriateness of the *neighbor* instrument but not the *non-neighbor* set with these data.[13]

---

[11] The parameter associated with secondary schools is one-third larger in the model that conditions on all four programs but it corresponds to a different policy experiment. In the schools only model, the secondary school parameter measures the marginal effect of secondary school provision conditional on fixed primary schools with the other two programs variable. In the model with all four programs, the secondary school parameter measures the marginal effect of secondary school provision conditional on all three other programs fixed. The spatial covariation of programs in conjunction with the impact of non-school programs on school enrollment can account for the different parameters in these two different conditional demand equations.
[12] Strictly speaking, the linear models that we estimate are *group* linear probability models since the outcome is the mean value for the sub-district of the binary outcome, as it is for the group probit. To simplify language, we drop the adjective group from this estimator.
[13] The robustness of results in Table 5 and Table 7 below were tested with a reduced instrument set that excluded parental age and natural shock variables (recent history of droughts, floods,

Column (1) of Table 6 uses the *neighbor* sub-district instruments and conditions on only school availability in the determination of boy's school enrollment. As before, column (2) adds family planning clinics and *PUSKESMAS* to the specification. Once more, Hansen's J-test fails to reject the null hypothesis in either specification at the 0.10 level (p=.15 and p=.14, respectively), providing some confidence in the validity of this instrument set, and the schools only specification is sufficient to not reject overidentification. Also in common with girl's schooling, augmenting the instrument set with *non-neighbors* permits the test of redundancy of the non-neighbors instrument set (rejected) and orthogonality of the neighbor instrument (not rejected), but unlike girl's schooling, the J-tests are not rejected with the larger set of instruments. A 1 percent point increase in secondary school availability is estimated to increase the school enrollment rate of boys ages 10-18 by 0.55 percentage points in the model with only schools, an effect that is seven times larger than the exogenous program model of Table 4, but smaller than the group probit effect (0.82).

Tables 7 and 8 estimate with GMM models of recent fertility and contraceptive use. Column (1) of each table estimates the effect of family planning programs on these two outcomes without conditioning on other programs, as in the usual policy function. In each case, the J-test rejects the null hypothesis (p=.008 and p=.000) decisively. [14] Adding *PUSKESMAS* public health clinics, which arguably directly affect maternal and child health, is sufficient to lead to non-rejection for fertility in column 2 of Table 7 (p=0.155) but not for contraceptive use in column 2 of Table 8 (p=.000). As noted above, the "failure" of these instruments may reflect their inclusion in the social planner's allocation of all programs, including schools, and the importance of those other programs in fertility and contraceptive choice. The result of column (3) of Table 7 and 8 are consistent with this view. Adding school programs to the specifications increases the *p*-values of the estimated J-tests (p=.20 and p=.22 for fertility and contraception, respectively), leading to the non-rejection of the null hypothesis. In particular, grade school availability has a statistically significant and negative effect on recent fertility and statistically significant and positive effect on contraceptive use, in common with the fertility/contraception literature. Adding school programs to the specification reduces the estimated effect of family planning clinics on contraceptive use by almost 85 percent, and renders the coefficient statistically insignificant with a standard error of nearly twice the coefficient. That is, conditional on public health clinics and schools, family

---

earthquakes and other disasters. Dropping these instruments ameliorates concerns that (i) parents who care more about their children's welfare may postpone having children until they are older and have more resources (parental age is endogenous), and (ii) shocks in other sub-districts may affect human capital in a particular neighboring sub-district through temporary migration (refugees) or variations in prices. Re-estimation with the reduced instrument (available on request) leaves the original results in Tables 5 and 7 essentially unaltered.

[14] The group probit functional form does not reject in the case of recent fertility (p=.105) but does in the case of contraceptive use (p=.000)

planning clinics have little discernible effect on contraceptive use, whereas schools and public health clinics have strong effects. In addition, the orthogonality of the *neighbor* instrument set is rejected (p=.000) in the model lacking the two school programs in column (5) of Table 8. In the model conditioning on all programs the null hypothesis of orthogonality of the *neighbor* instruments is not rejected (p=.76). We conclude that the statistically significant effect of family planning clinics in the models without schools (columns 1 and 2) apparently reflects the misspecification arising from invalid instruments when the full set of programs in not included.

For both recent fertility and contraceptive use and the full set of programs (column 6), augmenting the instrument set with *non-neighbors* leads to non-rejection of the J-test and orthogonality, rejection of redundancy, and roughly comparable estimates of program effects. The estimated effect of family planning clinics on fertility conditional on all programs in column(3) of Table 7 is 15 times larger than in the comparable specification without instruments in column (6) of Table 4. The estimates suggest that the expansion of *PUSKESMAS* and primary schools would importantly increase contraceptive use and decrease fertility.

## 6. Summary and conclusion

This paper proposes a novel instrumental variable method for program evaluation that only requires a single cross-section of data on the spatial intensity of programs and outcomes, and does not require the strong assumptions of fixed effects methods. The instruments are derived from a simple theoretical model of government decision-making that requires that the government's social welfare function is spatially weakly separable, that is, that the budgeting process is multi-stage with respect to administrative districts and sub-districts. Governments are assumed to be responsive to the attributes of places and their populations, rather than to the attributes of individuals, in making allocation decisions across space. If the attributes of a district rather than individual characteristics decide program placement, and the means (or higher moments) of outcomes for all "competing" districts enter into the government social welfare function, then the district means of individual and district exogenous determinants of these outcomes in *other* districts may be used as instruments for the placement of programs in any particular district. The assumption of weak separability of a social welfare function having as argument the means outcomes of every administrative unit (sub-district) is sufficient to generate *spatially decentralized budgeting*, The spatial instrumental variables model is then estimated and tested by GMM with a single cross-section of Indonesian census data.

The identification strategy proposed has broad applicability to the evaluation of public programs when the data consists of spatial variation in program intensity coupled with measures of outcomes for individuals who are matched to places. The assumptions of this model are not specific to Indonesia or developing countries more generally, and do not require an altruistic government. The model only requires that the attributes of places matter in the allocation of resources. Although it simplifies consideration of the spatial allocation process to think of it as a process in which larger administrative districts allocate a budget to smaller districts who in turn allocate them to even smaller administrative units, it not required that the government actually act in this manner, only that the central government use multi-stage budgeting as in the usual case of separable utility in consumer demand theory applied to a single economic agent.

The exclusion restrictions based upon multi-stage budgeting are that the sub-district means of individual-level variables and the environmental attributes of sub-districts in competing sub-districts influence program placement in a particular sub-district but not human capital outcomes in that sub-district conditional on program placement. Two sets of competing sub-districts are defined: (1) *neighbors*, which are means of the variables taken over the sub-districts that are spatially contiguous to sub-district ($\kappa,\ell$), and (2) *non-neighbors* , which are the means of the variables taken over the sub-districts that are in the same district (district $\ell$) as sub-district ($\kappa,\ell$) but are not spatially contiguous to sub-district ($\kappa,\ell$). We also investigate the simpler instrumental variable approach of Hausman and Taylor that only uses own-sub-district means of individual-level variables for identification. With spatial decentralization, the Hansen-Sargan J-tests fail to reject the null hypothesis of overidentification for the *neighbor* set of instrument in every specification (girl's and boy's schooling, recent fertility, and contraceptive use) containing all four government programs investigated (grade school, secondary school, public health clinics, and family planning clinics). The J-tests reject overidentification in every case with the Hausman-Taylor instruments. Adding the *non-neighbor* instrument set to the *neighbor* instrument set permits us to test for the orthogonality of the neighbor instrument set. These tests do not reject the null hypothesis in every case with the full set of programs. Program effects estimated with the neighbor instruments are very different in magnitude than the OLS estimates.

## References


Baum, Christopher, and Schaffer, Mark E. 2007. "Enhanced routines for instrumental variables/GMM estimation and testing," Boston College, Department of Economics, Working Paper No. 667, September.

Conley, Timothy, and E. Ligon. 2002. "Economic Distance and Cross-Country Spillovers." *Journal of Economic Growth* 7(2): 157-187.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review* 91(4): 795-813.

Hausman, Jerry A., and Taylor, William E. 1981. "Panel Data and Unobservable Individual Effects." *Econometrica* 49(6): 1377-1398.

Kleibergen, Frank, and Schaffer, Mark E. 2007. "RANKTEST: Stata module to testing the rank of a matrix using the Kleibergen-Paap rk statistic." Available at: http://ideas.repec.org/c/boc/bocode/s456865.html

Miguel, Edward, and Kremer, Michael. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72(1): 159-217.

Pitt, Mark M., Rosenzweig, Mark R., and M. N. Hassan. 1990. "Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries." *American Economic Review* 80: 1139-1156.

Pitt, Mark M., Rosenzweig, Mark R., and Gibbons, Donna M. 1993. "The Determinants and Consequences of the Placement of Government Programs in Indonesia." *World Bank Economic Review* 7(3): 319-348.

Rosenzweig, Mark R., and Wolpin, Kenneth I. 1986. "Evaluating the Effects of Optimally Distributed Public Programs: Child Health and Family Planning Interventions." *American Economic Review* 76(3): 470-482.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64(5): 416-424.

## Table 1: Means and standard deviations for the endogenous variables

| Variables | |
|---|---|
| ***Outcomes*** | |
| Current school enrollment for girls ages 10-18 years | 0.593<br>(0.196)<br>N=2921 |
| Current school enrollment for boys ages 10-18 years | 0.659<br>(0.178)<br>N=2919 |
| Whether last child's year of birth lies between 1978-1980 for women ages 21-30 years | 0.689<br>(0.163)<br>N=2914 |
| Whether any contraceptives are currently being used by women ages 21-30 years | 0.280<br>(0.244)<br>N=3033 |
| ***Programs*** | |
| Proportion of households in villages | |
| with grade schools | 0.774<br>(0.279)<br>N=2921 |
| with PUSKESMAS clinics | 0.245<br>(0.196)<br>N=2921 |
| with family planning clinics | 0.486<br>(0.335)<br>N=2921 |
| with junior or secondary schools | 0.394<br>(0.388)<br>N=2921 |

Standard deviations in parentheses. "N" denotes the number of *sub-district (kecamatan)* observations.

## Table 2: Means and standard deviations for the exogenous variables

| Variable | *Sub-district* Mean (1) | *Sub-district* SD (2) | *Neighboring* *Sub-districts* SD (3) | *Non-neighboring* Sub-districts SD (4) |
|---|---|---|---|---|
| ***Environmental variables*** | | | | |
| Proportion of households in villages | | | | |
| with urban status | 0.129 | 0.262 | 0.186 | 0.180 |
| with drought in the last five years | 0.263 | 0.284 | 0.207 | 0.167 |
| with flood in the last five years | 0.242 | 0.268 | 0.192 | 0.149 |
| with earthquake in the last five years | 0.090 | 0.213 | 0.182 | 0.158 |
| with other shocks in the last five years | 0.156 | 0.193 | 0.129 | 0.104 |
| with a coastal environment | 0.169 | 0.283 | 0.227 | 0.194 |
| ***Individual and household attributes*** | | | | |
| Dummy for household religion is Islam | 0.826 | 0.325 | 0.013 | 0.302 |
| Dummy for household religion is Christianity | 0.131 | 0.288 | 0.014 | 0.260 |
| Land owned by household (acres) | 0.648 | 0.718 | 8.620 | 5.115 |
| Square of land owned by household (acres) | 1.745 | 5.379 | 98.476 | 57.773 |
| Dummy for household owns its own home | 0.921 | 0.124 | 0.344 | 0.184 |
| Dummy for household head's language is Indonesian | 0.074 | 0.193 | 0.263 | 0.150 |
| Mother's age (years) | 40.308 | 2.722 | 0.560 | 3.398 |
| Household head's age (years) | 46.068 | 3.273 | 21.091 | 13.010 |
| Mother's schooling (years) | 2.441 | 1.611 | 15.688 | 3.466 |
| Household head's schooling (years) | 3.422 | 1.733 | 1.018 | 1.158 |
| Square of mother's schooling (years) | 15.173 | 14.498 | 13.703 | 8.983 |
| Square of household head's schooling (years) | 23.542 | 18.863 | 3.425 | 13.488 |
| Proportion of households in villages with urban status | | | | |
| interacted with land owned by household | 2.680 | 8.515 | 4.670 | 3.612 |
| interacted with square of land owned by household | 0.764 | 7.523 | 3.494 | 2.802 |
| interacted with dummy for household owns home | 0.101 | 0.194 | 0.132 | 0.123 |
| interacted with dummy for head’s lang. is Indonesian | 0.032 | 0.130 | 0.107 | 0.094 |
| interacted with mother’s schooling | 0.538 | 1.377 | 0.973 | 0.998 |
| interacted with household head’s schooling | 0.701 | 1.729 | 1.234 | 1.260 |

“SD” denotes standard deviation. Since the means for *neighboring* sub-districts (*kecamatans*) and the means for *non-neighboring* sub-districts are approximately the same as the sub-district means, the table presents only SDs in columns (3) and (4).

**Table 3: Models that use Hausman-Taylor instruments**

| | Current school enrollment for girls ages 10-18 | | Current school enrollment for boys ages 10-18 | | Whether last child's year of birth lies between 1978-1980 for women ages 21-30 | | Whether any contraceptives are currently being used by women ages 21-30 | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Proportion of households in villages | | | | | | | | |
| with grade schools | 0.009 | -0.020 | 0.084*** | 0.159** | --- | -0.278*** | --- | 0.825*** |
| | (0.028) | (0.059) | (0.026) | (0.063) | | (0.059) | | (0.142) |
| with junior or secondary schools | 0.175*** | 0.263*** | 0.187*** | 0.272*** | --- | 0.031 | --- | 0.033 |
| | (0.037) | (0.050) | (0.031) | (0.044) | | (0.046) | | (0.091) |
| with family planning clinics | --- | 0.093 | --- | -0.050 | -0.312*** | -0.098 | 0.374*** | -0.498*** |
| | | (0.067) | | (0.073) | (0.034) | (0.071) | (0.053) | (0.177) |
| with PUSKESMAS clinics | --- | -0.309*** | --- | -0.337*** | --- | 0.164** | --- | -0.172 |
| | | (0.084) | | (0.076) | | (0.083) | | (0.192) |
| Sargan-Hansen test for over-identification | 52.906 | 40.876 | 59.199 | 53.736 | 44.587 | 31.277 | 57.696 | 32.666 |
| | [0.000] | [0.004] | [0.000] | [0.000] | [0.005] | [0.052] | [0.000] | [0.037] |
| Observations (individuals) | 82891 | 82891 | 82889 | 82889 | 87655 | 87655 | 95372 | 95372 |

Spatially clustered standard errors in parentheses. *p*-values in square brackets. *** Denotes significance at the 1% level, ** denotes significance at the 5% level, and * denotes significance at the 10% level.

## Table 4: Models in which programs are treated exogenously (OLS)

| | Current school enrollment for girls ages 10-18 | | Current school enrollment for boys ages 10-18 | | Whether last child's year of birth lies between 1978-1980 for women ages 21-30 | | Whether any contraceptives are currently being used by women ages 21-30 | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Proportion of households in villages | | | | | | | | |
| with grade schools | 0.015 | 0.021 | 0.014 | 0.031 | --- | -0.093*** | --- | 0.202*** |
| | (0.020) | (0.021) | (0.018) | (0.020) | | (0.021) | | (0.035) |
| with junior or secondary schools | 0.082*** | 0.094*** | 0.080*** | 0.100*** | --- | 0.023* | --- | -0.023 |
| | (0.016) | (0.016) | (0.017) | (0.017) | | (0.014) | | (0.025) |
| with family planning clinics | --- | 0.008 | --- | -0.014 | -0.045*** | -0.019 | 0.098*** | 0.053** |
| | | (0.014) | | (0.012) | (0.016) | (0.016) | (0.024) | (0.027) |
| with PUSKESMAS clinics | --- | -0.060*** | --- | -0.081*** | --- | -0.045** | --- | -0.074** |
| | | (0.017) | | (0.017) | | (0.022) | | (0.031) |
| Observations (individuals) | 82891 | 82891 | 82889 | 82889 | 87655 | 87655 | 95372 | 95372 |

Spatially clustered standard errors in parentheses. *** Denotes significance at the 1% level, ** denotes significance at the 5% level, and * denotes significance at the 10% level.

## Table 5: Current school enrollment for girls ages 10-18

| | GMM linear (1) | GMM linear (2) | GMM linear (3) | GMM linear (4) | GMM group probit (5) | GMM group probit (6) |
|---|---|---|---|---|---|---|
| Proportion of households in villages | | | | | | |
| with grade schools | 0.023 | 0.041 | 0.042 | 0.057 | 0.157 | 0.182 |
| | (0.074) | (0.081) | (0.039) | (0.045) | (0.264) | (0.300) |
| with junior or secondary schools | 0.621*** | 0.847*** | 0.361*** | 0.454*** | 2.241*** | 3.069*** |
| | (0.150) | (0.186) | (0.083) | (0.091) | (0.642) | (0.766) |
| with family planning clinics | --- | -0.004 | --- | 0.010 | --- | -0.048 |
| | | (0.065) | | (0.042) | | (0.239) |
| with PUSKESMAS clinics | --- | -0.251 | --- | -0.155** | --- | -0.978* |
| | | (0.154) | | (0.077) | | (0.598) |
| Neighboring sub-districts | Yes | Yes | Yes | Yes | Yes | Yes |
| Non-neighboring sub-districts | No | No | Yes | Yes | No | No |
| Hansen's *J* test | 26.327 | 21.381 | 63.467 | 62.201 | 22.266 | 17.075 |
| | [0.121] | [0.210] | [0.011] | [0.008] | [0.271] | [0.449] |
| Orthogonality of neighboring sub-districts | --- | --- | 19.108 [0.578] | 21.598 [0.423] | --- | --- |
| Redundancy of non-neighboring sub-districts | --- | --- | 92.836 [0.000] | 165.181 [0.000] | --- | --- |
| Observations (sub-districts) | 2921 | 2921 | 2921 | 2921 | 2921 | 2921 |

Spatially clustered standard errors in parentheses. *p*-values in square brackets. *** Denotes significance at the 1% level, ** denotes significance at the 5% level, and * denotes significance at the 10% level.

## Table 6: Current school enrollment for boys ages 10-18

| | GMM linear (1) | GMM linear (2) | GMM linear (3) | GMM linear (4) | GMM group probit (5) | GMM group probit (6) |
|---|---|---|---|---|---|---|
| Proportion of households in villages | | | | | | |
| with grade schools | 0.012 | 0.067 | 0.006 | 0.004 | 0.063 | 0.297 |
| | (0.065) | (0.076) | (0.040) | (0.050) | (0.251) | (0.289) |
| with junior or secondary schools | 0.552*** | 0.674*** | 0.359*** | 0.431*** | 2.229*** | 2.563*** |
| | (0.167) | (0.203) | (0.074) | (0.090) | (0.674) | (0.784) |
| with family planning clinics | --- | -0.057 | --- | 0.029 | --- | -0.169 |
| | | (0.058) | | (0.043) | | (0.213) |
| with PUSKESMAS clinics | --- | -0.059 | --- | -0.088 | --- | -0.281 |
| | | (0.113) | | (0.074) | | (0.369) |
| Neighboring sub-districts | Yes | Yes | Yes | Yes | Yes | Yes |
| Non-neighboring sub-districts | No | No | Yes | Yes | No | No |
| Hansen's *J* test | 25.398 | 21.168 | 48.312 | 47.662 | 23.691 | 21.599 |
| | [0.148] | [0.144] | [0.172] | [0.135] | [0.208] | [0.201] |
| Orthogonality of neighboring sub-districts | --- | --- | 22.088 | 25.634 | --- | --- |
| | | | [0.394] | [0.221] | | |
| Redundancy of non-neighboring sub-districts | --- | --- | 89.478 | 157.571 | --- | --- |
| | | | [0.000] | [0.000] | | |
| Observations (sub-districts) | 2919 | 2919 | 2919 | 2919 | 2919 | 2919 |

Spatially clustered standard errors in parentheses. *p*-values in square brackets. *** Denotes significance at the 1% level, ** denotes significance at the 5% level, and * denotes significance at the 10% level.

## Table 7: Whether last child's year of birth lies between 1978-1980 for women ages 21-30

| | GMM linear (1) | GMM linear (2) | GMM linear (3) | GMM linear (4) | GMM linear (5) | GMM linear (6) | GMM group probit (7) | GMM group probit (8) | GMM group probit (9) |
|---|---|---|---|---|---|---|---|---|---|
| Proportion of households in villages | | | | | | | | | |
| with grade schools | --- | --- | -0.186** | --- | --- | -0.062 | --- | --- | -0.696** |
| | | | (0.085) | | | (0.053) | | | (0.304) |
| with junior or secondary schools | --- | --- | -0.063 | --- | --- | 0.122 | --- | --- | -0.159 |
| | | | (0.184) | | | (0.096) | | | (0.623) |
| with family planning clinics | -0.216*** | -0.358*** | -0.269*** | -0.173*** | -0.265*** | -0.227*** | -0.712*** | -1.128*** | -0.848*** |
| | (0.043) | (0.077) | (0.086) | (0.029) | (0.050) | (0.057) | (0.165) | (0.255) | (0.316) |
| with PUSKESMAS clinics | --- | 0.289** | 0.319** | --- | 0.199** | 0.163* | --- | 0.849** | 1.007* |
| | | (0.135) | (0.151) | | (0.090) | (0.091) | | (0.411) | (0.530) |
| Neighboring sub-districts | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Non-neighboring sub-districts | No | No | No | Yes | Yes | Yes | No | No | No |
| Hansen's *J* test | 38.436 | 25.172 | 21.510 | 66.202 | 49.247 | 47.655 | 28.184 | 20.660 | 15.987 |
| | [0.008] | [0.155] | [0.204] | [0.008] | [0.150] | [0.136] | [0.105] | [0.356] | [0.525] |
| Orthogonality of neighboring sub-districts | --- | --- | --- | 14.793 | 13.094 | 11.572 | --- | --- | --- |
| | | | | [0.833] | [0.905] | [0.951] | | | |
| Redundancy of non-neighboring sub-districts | --- | --- | --- | 79.919 | 106.962 | 165.562 | --- | --- | --- |
| | | | | [0.000] | [0.000] | [0.000] | | | |
| Observations (sub-districts) | 2914 | 2914 | 2914 | 2914 | 2914 | 2914 | 2914 | 2914 | 2914 |

Spatially clustered standard errors in parentheses. *p*-values in square brackets. *** Denotes significance at the 1% level, ** denotes significance at the 5% level, and * denotes significance at the 10% level.

## Table 8: Whether any contraceptives are currently being used by women ages 21-30

| | GMM linear (1) | GMM linear (2) | GMM linear (3) | GMM linear (4) | GMM linear (5) | GMM linear (6) | GMM group probit (7) | GMM group probit (8) | GMM group probit (9) |
|---|---|---|---|---|---|---|---|---|---|
| Proportion of households in villages | | | | | | | | | |
| with grade schools | --- | --- | 0.559*** | --- | --- | 0.509*** | --- | --- | 2.673*** |
| | | | (0.122) | | | (0.079) | | | (0.456) |
| with junior or secondary schools | --- | --- | 0.594* | --- | --- | 0.276* | --- | --- | 2.219* |
| | | | (0.314) | | | (0.168) | | | (1.163) |
| with family planning clinics | 0.133** | 0.423*** | 0.068 | 0.119*** | 0.394*** | 0.025 | 0.790*** | 1.706*** | -0.051 |
| | (0.059) | (0.094) | (0.128) | (0.045) | (0.068) | (0.086) | (0.227) | (0.337) | (0.453) |
| with PUSKESMAS clinics | --- | -0.609*** | -0.630*** | --- | -0.598*** | -0.490*** | --- | -2.079*** | -2.184*** |
| | | (0.139) | (0.177) | | (0.113) | (0.120) | | (0.509) | (0.672) |
| Neighboring sub-districts | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Non-neighboring sub-districts | No | No | No | Yes | Yes | Yes | No | No | No |
| Hansen's $J$ test | 74.616 | 47.049 | 21.265 | 104.410 | 70.744 | 48.314 | 62.383 | 47.145 | 15.640 |
| | [0.000] | [0.000] | [0.215] | [0.000] | [0.002] | [0.122] | [0.000] | [0.000] | [0.550] |
| Orthogonality of neighboring sub-districts | --- | --- | --- | 24.964 | 99.065 | 16.161 | --- | --- | --- |
| | | | | [0.249] | [0.000] | [0.761] | | | |
| Redundancy of non-neighboring sub-districts | --- | --- | --- | 65.258 | 33.167 | 158.757 | --- | --- | --- |
| | | | | [0.000] | [0.044] | [0.000] | | | |
| Observations (sub-districts) | 3033 | 3033 | 3033 | 3033 | 3033 | 3033 | 3033 | 3033 | 3033 |

Spatially clustered standard errors in parentheses. $p$-values in square brackets. *** Denotes significance at the 1% level, ** denotes significance at the 5% level, and * denotes significance at the 10% level.