The Harmony of Programs Package: Quasi-experimental Evidence on Health and Nutrition Interventions in Rural Senegal

Théophile T. Azomahou[a], Fatoumata L. Diallo[b], Wladimir Raymond[c]

(a) UNU-MERIT and Maastricht University
(b) University Cheikh-Anta-Diop and CRES
(c) STATEC Luxembourg

UNU-WIDER Development Conference: Human Capital and Growth
Helsinki, 6-7 June 2016

Outline

Motivation

Why malnutrion and intestinals parasites are important issues in developing countries?

- In underdeveloped countries, hundreds of millions of children suffer from poverty, health, morbidity and malnutrition.
- Severe malnutrition can cause delays or even deficits in cognitive development.
- Intestinal worms are endemic in tropical and subtropical regions.
- At global level, because of its negative impact on health and education, malnutrition contributes to weaken human capital accumulation:
  - early growth faltering (exposure in utero)
  - nutritional effects on younger children (0-2 years)
  - slower brain development and effects on the delay in school

Motivation (cont'd)

Importance of evaluating nutrition and health programs

- Programme evaluation has become an important tool to inform policy makers about the efficient allocation of resources and for the improvement of existing policies.
- In Senegal, poverty and vulnerability are higher in rural areas.
- The government of Senegal who supports since over 10 years nutritional and health programs in rural schools always tries to know to what extent these programs produce positives results.

Motivation (cont'd)

Empirical studies do not draw the same conclusions

School meals and school performance: some contributions

- **No evidence**: Ahmed (2004), Kazianga et al.(2009), Tan et al. (1999), Simeon and Grantham-McGregor (1989).
- **Positive effect**: Vermeerrch and Kremer (2004), Cueto and Chinen (2007), Ahmed (2004), Simeon and Grantham-McGregor (1989).
- **Negative effect**: Ahmed and del Ninno (2002), Ahmed (2004).

Motivation (cont'd)

Empirical studies do not draw the same conclusions

Deworming and school performance: some contributions

- **No evidence**: Miguel and Kremer (2004), Kvalsig et al. (1991), Nokes et al. (1992).
- **Positive effect**: Kvalsig et al. (1991).
- **Negative effect**: Miguel and Kremer (2004).
- **Home vs. school deworming**: Azomahou and Diallo (2016). Deworming at school has a positive effect on pupils' performance while deworming at home has a negative impact. This result indicates that the use of widely spread traditional deworming medicines should be discouraged.
- There is no study that try to measure both the impact of a lunch and a deworming program at school while estimating the determinants of school performance in a joint framework.

Motivation (cont'd)

Aim of this study

- Assess the impact of school deworming and meal as programs package on test scores, enrollment, promotion and dropout rate while elaborating on the determinants of school performance.

- Contributions of the paper:
  - i) A new dataset (quasi-experimental)
  - ii) Advantages of package programs: Low cost and more effective than simple programs
  - iii) Econometric framework: Double-index selection models (double endogenous selection vs. generalized Roy)
  - iv) Wide range of treatment effects
  - v) Policy analysis: a) cost-effectiveness, b) welfare benefit of programs

Data and variables

Data

- School deworming and meal programs implemented in early 2000 by the World Food Programme (WFP) and the Government of Senegal
- Primary data collected by the 'Consortium pour la recherche Economique et Social (CRES)' and the Ministry of Education as part of an experimental program on school canteens and deworming in rural Senegal.
- New and rich data set: an important amount of work (cleaning, recoding and imputing) has been done to make it useable.
- Sample of about 4500 pupils for 160 schools.
- Data provide information about pupils, schools, households and communities characteristics.

Data and variables (cont'd)

Variables

- Outcome indicators:
  i) Scores: aggregate, French and Math
  ii) Enrollment, promotion and dropout rate

  ! **Debate on the relevance of such outcomes. Ideally, one would like to have pure nutritional outcomes (e.g. child growth, etc).**

- Control variable gathered into four categories:
  i) Pupils characteristics (gender, age, Islamic school,...)
  ii) Schools characteristics (distance to school, class size, water point,...)
  iii) Household characteristics (food, health and education expenditures,...)
  iv) Community environment (college, children labor, domestic chores,...)
- Double treatment (T_1: Dworm=1, T_2: Dmeal=1)
- **Full list of variables: see paper**

Treatment status: Deworming () vs. meal

		Meal program T_2)		Total (margins for T_1)
		0	1	
Deworming program (T_1)	0	$\boxed{65.013\%}$	22.914%	87.927%
		(0.476)	(0.420)	
	1	8.202%	3.871%	12.073%
		(0.274)	(0.192)	
Total (margins for T_2)		73.215%	26.785%	100%

Aggregate score: by treatment status

		Meal program	
		0	1
Deworming program	0	37.687	41.771
		(19.306)	(18.992)
	1	36.694	$\boxed{47.663}$
		(16.561)	(14.266)

French score: by treatment status

		Meal program	
		0	1
Deworming program	0	38.413	40.742
		(21.071)	(21.028)
	1	35.366	$\boxed{45.242}$
		(19.817)	(17.885)

Math score: by treatment status

		Meal program	
		0	1
Deworming program	0	36.965	42.678
		(21.121)	(21.204)
	1	37.627	$\boxed{50.084}$
		(17.687)	(16.699)

Enrollment rate: by treatment status

		Meal program	
		0	1
Deworming program	0	-31.404	$\boxed{7.631}$
		(60.005)	(37.574)
	1	-32.281	-20.358
		(34.715)	(51.624)

Promotion rate: by treatment status

		Meal program	
		0	1
Deworming program	0	79.152	78.760
		(12.959)	(11.740)
	1	73.345	$\boxed{81.887}$
		(14.798)	(6.760)

Dropout rate: by treatment status

		Meal program	
		0	1
Deworming program	0	16.603	15.072
		(12.989)	(9.302)
	1	15.181	$\boxed{10.191}$
		(12.400)	(0.950)

DES model (cont'd)

Specification

Probit selection equations: Let T_{1i}^* and T_{2i}^* denote two *latent* (unobserved) variables, which are assumed to be functions of *observed* characteristics $\mathbf{w}_{ji}$ ($j = 1$ or 2) of N households/firms ($i = 1, \cdots, N$). Formally,

$$T_{1i}^* = \gamma_1' \mathbf{w}_{1i} + \mu_{1i}, \qquad (1)$$

$$T_{2i}^* = \gamma_2' \mathbf{w}_{2i} + \mu_{2i}, \qquad (2)$$

where γ_j denotes the vectors of parameters to be estimated, and μ_{ji} denotes the usual error terms.

The observed counterparts to T_{1i}^* and T_{2i}^*, denoted by T_{1i} and T_{2i}, are defined as

$$T_{1i} = \mathbf{1}[T_{1i}^* > 0], \qquad (3)$$

$$T_{2i} = \mathbf{1}[T_{2i}^* > 0], \qquad (4)$$

where $\mathbf{1}[\cdot]$ denotes the indicator function.

DES model (cont'd)

Specification

Outcome equation: The outcome for individual i, y_i, is given by

$$y_i = \boldsymbol{\beta}' \mathbf{x}_i + \delta_1 T_{1i} + \delta_2 T_{2i} + \theta T_{1i} T_{2i} + \varepsilon_i, \tag{5}$$

where $\mathbf{x}_i$ denotes control variables (e.g. household income, etc.), $\boldsymbol{\beta}$, δ_j and θ are parameter vectors and scalars to be estimated, and ε_i denotes the error term.

By including the interaction term, $T_{1i} T_{2i}$, as a regressor in equation (5), we can isolate the exclusive effect of either treatment and their joint effect, while estimating *complementarity* ($\theta > 0$) or *substitutability* ($\theta < 0$) of policies T_{1i}, T_{2i}.

Equations (1)-(5) is a generalization of the dummy endogenous variable model of Heckman (1978).

DES model (cont'd)

Estimations

- **FIML (a)** method: Assumptions

  We make the following distributional assumption: $(\mu_{1i}, \mu_{2i}, \varepsilon_i)'$ is normally distributed with vector mean $\mathbf{0}$ and covariance matrix:

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & & \\ \rho_{\mu_1\mu_2} & 1 & \\ \rho_{\mu_1\varepsilon}\sigma_\varepsilon & \rho_{\mu_2\varepsilon}\sigma_\varepsilon & \sigma_\varepsilon^2 \end{pmatrix}$$

  The likelihood function of the model consists of four parts following from the contributions of the two selections: $(T_{1i} = 1,\, T_{2i} = 1)$, $(T_{1i} = 1,\, T_{2i} = 0)$, $(T_{1i} = 0,\, T_{2i} = 1)$, $(T_{1i} = 0,\, T_{2i} = 0)$

DES model (con'd)

- **FIML (b)**: The likelihood. Combination of the four contributions:

$$
\begin{aligned}
L = \prod_{i=1}^{N} & \left[\int_{-\gamma_1' \mathbf{w}_{1i}}^{\infty} \int_{-\gamma_2' \mathbf{w}_{2i}}^{\infty} K(y_i, \mathbf{x}_i, \cdot) f_2(\mu_{1i}, \mu_{2i} | y_i) f_1(y_i) d\mu_{1i} d\mu_{2i} \right]^{T_{1i} T_{2i}} \\
& \left[\int_{-\gamma_1' \mathbf{w}_{1i}}^{\infty} \int_{-\infty}^{-\gamma_2' \mathbf{w}_{2i}} K(y_i, \mathbf{x}_i, \cdot) f_2(\mu_{1i}, \mu_{2i} | y_i) f_1(y_i) d\mu_{1i} d\mu_{2i} \right]^{T_{1i}(1-T_{2i})} \\
& \left[\int_{-\infty}^{-\gamma_1' \mathbf{w}_{1i}} \int_{-\gamma_2' \mathbf{w}_{2i}}^{\infty} K(y_i, \mathbf{x}_i, \cdot) f_2(\mu_{1i}, \mu_{2i} | y_i) f_1(y_i) d\mu_{1i} d\mu_{2i} \right]^{(1-T_{1i}) T_{2i}} \\
& \left[\int_{-\infty}^{-\gamma_1' \mathbf{w}_{1i}} \int_{-\infty}^{-\gamma_2' \mathbf{w}_{2i}} K(y_i, \mathbf{x}_i, \cdot) f_2(\mu_{1i}, \mu_{2i} | y_i) f_1(y_i) d\mu_{1i} d\mu_{2i} \right]^{(1-T_{1i})(1-T_{2i})}
\end{aligned}
\tag{6}
$$

where $K(y_i, \mathbf{x}_i, \cdot) = \frac{1}{\sigma_\varepsilon} \phi_1 \left(\frac{y_i - \beta' \mathrm{x}_i - A_i(T_{1i}, T_{2i})}{\sigma_\varepsilon} \right)$, ϕ_1 denotes the univariate standard normal density function and $A_i(T_{1i}, T_{2i}) \equiv \delta_1 T_{1i} + \delta_2 T_{2i} + \theta T_{1i} T_{2i}$.

Difficulty: evaluate the double integrals in (6) as bivariate CDFs conditional on a third random variable, y_i.

DES model (cont'd)

Estimations

- **Two step method (cont'd)**: Assumptions

  The regression (5) is the equation of interest. Relying on Blundell and Costa Dias (2000, 2008), the population regression can be written in the form of a conditional expectation, i.e.

  $$E(y_i | T_{1i}, T_{2i}, \mathbf{x}_i) = \boldsymbol{\beta}' \mathbf{x}_i + \delta_1 T_{1i} + \delta_2 T_{2i} + \theta T_{1i} T_{2i} + E(\varepsilon_i | T_{1i}, T_{2i}, \mathbf{x}_i). \quad (7)$$

  Since T_{1i} and T_{2i} are endogenous, $E(\varepsilon_i | T_{1i}, T_{2i}, \mathbf{x}_i) \neq 0$ and the (OLS) estimator of $\boldsymbol{\beta}$, δ_1, δ_2 and θ is inconsistent.

  The endogeneity of T_{ji} ($j = 1$ or 2) comes from the fact that T_{ji} depends on μ_{ji} and the latter is correlated with ε_i. Hence, the endogeneity is accounted for by taking the correlations $\rho_{\mu_1 \varepsilon}$ and $\rho_{\mu_2 \varepsilon}$ into account.

DES model (cont'd)

- **Two step method (cont'd)**: Using the definition of T_{1i} and T_{2i} and the latent Eqs. (1) and (2):

$$\begin{aligned} E(\varepsilon_i | T_{1i}, T_{2i}, \mathbf{x}_i) = {} & T_{1i} T_{2i} \underbrace{E(\varepsilon_i | \mu_{1i} > -\gamma_1' \mathbf{w}_{1i}, \mu_{2i} > -\gamma_2' \mathbf{w}_{2i}, \mathbf{x}_i)}_{E(\varepsilon_i|>,>)} \\ & + T_{1i}(1 - T_{2i}) \underbrace{E(\varepsilon_i | \mu_{1i} > -\gamma_1' \mathbf{w}_{1i}, \mu_{2i} \leq -\gamma_2' \mathbf{w}_{2i}, \mathbf{x}_i)}_{E(\varepsilon_i|>,\leq)} \\ & + (1 - T_{1i}) T_{2i} \underbrace{E(\varepsilon_i | \mu_{1i} \leq -\gamma_1' \mathbf{w}_{1i}, \mu_{2i} > -\gamma_2' \mathbf{w}_{2i}, \mathbf{x}_i)}_{E(\varepsilon_i|\leq,>)} \\ & + (1 - T_{1i})(1 - T_{2i}) \underbrace{E(\varepsilon_i | \mu_{1i} \leq -\gamma_1' \mathbf{w}_{1i}, \mu_{2i} \leq -\gamma_2' \mathbf{w}_{2i}, \mathbf{x}_i)}_{E(\varepsilon_i|\leq,\leq)} . \end{aligned} \tag{8}$$

The conditional expectations in Eq. (8) involve the truncated trivariate normal distribution of the form $E(x|y > a, z > b), \cdots, E(x|y \leq \cdot, z \leq \cdot)$.

DES model (con'd)

- **Two step method (con'd)**: Using the moment-generating function formula along the lines of Muthén (1990), these expectations are shown to be (see Appendix for details):

$$E(\varepsilon_i| >, >) = \frac{\sigma_\varepsilon \rho_{\mu_1 \varepsilon} \phi_1(\gamma_1' \mathbf{w}_{1i})}{\Phi_2\left(\gamma_1' \mathbf{w}_{1i}, \gamma_2' \mathbf{w}_{2i}, \rho_{\mu_1 \mu_2}\right)} \Phi_1 \left(\frac{\gamma_2' \mathbf{w}_{2i} - \rho_{\mu_1 \mu_2} \gamma_1' \mathbf{w}_{1i}}{\sqrt{1 - \rho^2_{\mu_1 \mu_2}}} \right) + \frac{\sigma_\varepsilon \rho_{\mu_2 \varepsilon} \phi_1(\gamma_2' \mathbf{w}_{2i})}{\Phi_2\left(\gamma_1' \mathbf{w}_{1i}, \gamma_2' \mathbf{w}_{2i}, \rho_{\mu_1 \mu_2}\right)} \Phi_1 \left(\frac{\gamma_1' \mathbf{w}_{1i} - \rho_{\mu_1 \mu_2} \gamma_2' \mathbf{w}_{2i}}{\sqrt{1 - \rho^2_{\mu_1 \mu_2}}} \right)$$

till ....

$$E(\varepsilon_i| \leq, \leq) = -\frac{\sigma_\varepsilon \rho_{\mu_1 \varepsilon} \phi_1(\gamma_1' \mathbf{w}_{1i})}{\Phi_2\left(-\gamma_1' \mathbf{w}_{1i}, -\gamma_2' \mathbf{w}_{2i}, \rho_{\mu_1 \mu_2}\right)} \Phi_1 \left(\frac{\rho_{\mu_1 \mu_2} \gamma_1' \mathbf{w}_{1i} - \gamma_2' \mathbf{w}_{2i}}{\sqrt{1 - \rho^2_{\mu_1 \mu_2}}} \right) - \frac{\sigma_\varepsilon \rho_{\mu_2 \varepsilon} \phi_1(\gamma_2' \mathbf{w}_{2i})}{\Phi_2\left(-\gamma_1' \mathbf{w}_{1i}, -\gamma_2' \mathbf{w}_{2i}, \rho_{\mu_1 \mu_2}\right)} \Phi_1 \left(\frac{\rho_{\mu_1 \mu_2} \gamma_2' \mathbf{w}_{2i} - \gamma_1' \mathbf{w}_{1i}}{\sqrt{1 - \rho^2_{\mu_1 \mu_2}}} \right)$$

DES model (cont'd)

- **Two step method (cont'd)**: Replacing the expressions, the conditional expectations of equation (8) yields after factorization

$$\begin{aligned} y_i &= \boldsymbol{\beta}'\mathbf{x}_i + \delta_1 T_{1i} + \delta_2 T_{2i} + \theta T_{1i} T_{2i} \\ &+ \underbrace{\sigma_\varepsilon \rho_{\mu_1 \varepsilon}}_{\eta_1} h_1(T_{1i}, T_{2i}) + \underbrace{\sigma_\varepsilon \rho_{\mu_2 \varepsilon}}_{\eta_2} h_2(T_{1i}, T_{2i}) + \nu_i \end{aligned} \tag{9}$$

where $E[\nu_i | \mathbf{x}_i, h_1(T_{1i}, T_{2i}), h_2(T_{1i}, T_{2i})] = 0$ and η_1 and η_2 are additional parameters to be estimated, with:

$$\begin{aligned} h_1(T_{1i}, T_{2i}) &= \lambda_1^{++} T_{1i} T_{2i} + \lambda_1^{+-} T_{1i}(1 - T_{2i}) \\ &- \lambda_1^{-+}(1 - T_{1i}) T_{2i} - \lambda_1^{--}(1 - T_{1i})(1 - T_{2i}) \end{aligned}$$

$$\begin{aligned} h_2(T_{1i}, T_{2i}) &= \lambda_2^{++} T_{1i} T_{2i} - \lambda_2^{+-} T_{1i}(1 - T_{2i}) \\ &+ \lambda_2^{-+}(1 - T_{1i}) T_{2i} - \lambda_2^{--}(1 - T_{1i})(1 - T_{2i}) \end{aligned}$$

and the λ's are generalizations of the inverse Mill's ratio.

DES model (cont'd)

- **Two step method (cont'd)**: Implementation

  In practice, one problem occurs in that $h_1(T_{1i}, T_{2i})$ and $h_2(T_{1i}, T_{2i})$ are unobserved as they are functions of the unobserved parameters γ_1, γ_2 and $\rho_{\mu_1\mu_2}$, hence the two-step approach:

  1. Obtain consistent and efficient (under normality) estimates for γ_1, γ_2 and $\rho_{\mu_1\mu_2}$ by estimating a bivariate probit. Compute $\widehat{h}_1(T_{1i}, T_{2i})$ and $\widehat{h}_2(T_{1i}, T_{2i})$ by estimating the different λ's using $\widehat{\gamma}_1$, $\widehat{\gamma}_2$ and $\widehat{\rho}_{\mu_1\mu_2}$.
  2. Use $\widehat{h}_1(T_{1i}, T_{2i})$ and $\widehat{h}_2(T_{1i}, T_{2i})$ as additional regressors in Eq. (9) alongside $\mathbf{x}_i$, T_{1i} and T_{2i} and apply OLS to equation (9). Since we use their estimates in lieu of $h_1(T_{1i}, T_{2i})$ and $h_2(T_{1i}, T_{2i})$, the conventional standard errors are not valid and need to be corrected by generalizing the results of Heckman (1976,1979), or by using techniques of simulation or bootstrap.

Roy model (cont'd)

Specification

We suggest a generalization of Roy's model also known as the endogenous switching regression model with four regimes. The selection equations are the same as (1)-(4). Let us define the four corresponding outcomes as y_{i11}, y_{i10}, y_{i01} and y_{i00} respectively:

$$y_{i11} = \beta'_{11}\mathbf{x}_i + \varepsilon_{i11}, \tag{10a}$$
$$y_{i10} = \beta'_{10}\mathbf{x}_i + \varepsilon_{i10}, \tag{10b}$$
$$y_{i01} = \beta'_{01}\mathbf{x}_i + \varepsilon_{i01}, \tag{10c}$$
$$y_{i00} = \beta'_{i00}\mathbf{x}_i + \varepsilon_{i00}. \tag{10d}$$

Furthermore, y_{i11}, y_{i10}, y_{i01} and y_{i00} are not jointly observed, but only one of these outcomes is observed at a time.

Roy model (cont'd)

Estimations

- **ML (a)** method: Assumptions

  We maintain the normality assumption. In this case, $(\mu_{1i}, \mu_{2i}, \varepsilon_{i11}, \varepsilon_{i10}, \varepsilon_{i01}, \varepsilon_{i00})'$ is normally distributed with vector mean **0** and covariance matrix $\mathbf{\Omega}$ defined as

$$\mathbf{\Omega} = \begin{pmatrix} 1 & & & & & \\ \rho_{\mu_1\mu_2} & 1 & & & & \\ \rho_{\mu_1\varepsilon_{11}}\sigma_{\varepsilon_{11}} & \rho_{\mu_2\varepsilon_{11}}\sigma_{\varepsilon_{11}} & \sigma^2_{\varepsilon_{11}} & & & \\ \rho_{\mu_1\varepsilon_{10}}\sigma_{\varepsilon_{10}} & \rho_{\mu_2\varepsilon_{10}}\sigma_{\varepsilon_{10}} & 0 & \sigma^2_{\varepsilon_{10}} & & \\ \rho_{\mu_1\varepsilon_{01}}\sigma_{\varepsilon_{01}} & \rho_{\mu_2\varepsilon_{01}}\sigma_{\varepsilon_{01}} & 0 & 0 & \sigma^2_{\varepsilon_{01}} & \\ \rho_{\mu_1\varepsilon_{00}}\sigma_{\varepsilon_{00}} & \rho_{\mu_2\varepsilon_{00}}\sigma_{\varepsilon_{00}} & 0 & 0 & 0 & \sigma^2_{\varepsilon_{00}} \end{pmatrix}$$

Roy model

- **Two step (b)**: Procedure

Using results of the previous conditional expectations, we write the regression of the sub-populations as

$$y_{i11} = \beta'_{11}\mathbf{x}_i + \sigma_{\varepsilon_{11}}\rho_{\mu_1\varepsilon_{11}}\lambda_1^{++} + \sigma_{\varepsilon_{11}}\rho_{\mu_2\varepsilon_{11}}\lambda_2^{++} + \nu_{i11}, \tag{11a}$$

$$y_{i10} = \beta'_{10}\mathbf{x}_i + \sigma_{\varepsilon_{10}}\rho_{\mu_1\varepsilon_{10}}\lambda_1^{+-} - \sigma_{\varepsilon_{10}}\rho_{\mu_2\varepsilon_{10}}\lambda_2^{+-} + \nu_{i10}, \tag{11b}$$

$$y_{i01} = \beta'_{01}\mathbf{x}_i - \sigma_{\varepsilon_{01}}\rho_{\mu_1\varepsilon_{01}}\lambda_1^{-+} + \sigma_{\varepsilon_{01}}\rho_{\mu_2\varepsilon_{01}}\lambda_2^{-+} + \nu_{i01}, \tag{11c}$$

$$y_{i00} = \beta'_{00}\mathbf{x}_i - \sigma_{\varepsilon_{00}}\rho_{\mu_1\varepsilon_{00}}\lambda_1^{--} - \sigma_{\varepsilon_{00}}\rho_{\mu_2\varepsilon_{00}}\lambda_2^{--} + \nu_{i00}, \tag{11d}$$

with $E(\nu_{i11}|\mathbf{x}_i, \lambda_1^{++}, \lambda_2^{++}) = ... = E(\nu_{i00}|\mathbf{x}_i, \lambda_1^{--}, \lambda_2^{--}) = 0$, and where the expressions of the λ's are the same as previous.

Treatment effects

Treatment Effects

- **Exclusive effects** of T_1 resp. T_2: the marginal effect of T_1 resp. T_2 conditional on the fact that agents are not in the alternative programme. Such effects allow to measure the impact T_1 or T_2 only on the outcome y, given controls $\mathbf{x}$.
- **Global effect**: is the effect of both programmes taken together.
- **Additional effect**: effect following from having additionally another programme. It is given by the difference between the global effect and the exclusive effects.
- **Relative effects**:

  i) Effect of (T_1, T_2) whole package vs. T_1.
  ii) Effect of (T_1, T_2) whole package vs. T_2.
  iii) Effect of T_1 vs. T_2.
- **Sequential effects**:

  i) Sequence (T_1, T_2)
  i) Sequence (T_2, T_1)
- **Substitution effects**

Summary of findings (cont'd)

Specification check

: DISM[a] model: FIML[b]

Correlation	Aggregate score Coef.	Aggregate score Std. Err.	French score Coef.	French score Std. Err.	Math score Coef.	Math score Std. Err.
$\rho 12$	-0.071^{*}	0.039	-0.066^{*}	0.039	-0.067^{*}	0.039
$\rho 13$	0.567^{***}	0.058	0.494^{***}	0.070	0.528^{***}	0.060
$\rho 23$	0.098	0.068	-0.010	0.068	0.181^{***}	0.063
# Obs.			3487			

Notes: [a]Double-Index Selection Model; [b]Full Information Maximum Likelihood Estimation.
Significance levels: $*$ 10% $**$ 5% $***$ 1%

Remark: 13 free parameters to estimate for the Roy model. FIML not feasible due to few obs. to reach convergence, so rely only on Heckman two steps.

Summary of findings (cont'd)

Complementary vs. substitutability

The two programs are complementary in the goal of increasing scores and promotion rates. Contrariwise, they are substitutes with the aim of reducing dropouts.

Treatment effects

- Score outcomes
  - i) Positive and significant additional, exclusive and global average treatment effects (ATE).
  - ii) The impact of the meals program on the scores is greater than that the deworming program.
  - iii) The combination of the two programs (package) has a greater impact. This result reinforces the complementary finding. Moreover, the relative effect of the package vs. the deworming alone is greater than that compared to the canteen only.
  - iv) ATET: the exclusive, global, and additional effects are positive and significant. It should be noted that the effects on the treated are larger than the ATE.

Summary of findings (cont'd)

Treatment effects (con't)

- Score outcomes:

  v) ATENT: the exclusive effect of deworming is negative while the effect of canteen is positive. The combination of the two programs greatly increases scores.

  vi) The sequential effects indicate that for the pupils in the treated group, the impact of the package performs better if the school meals is introduced before deworming. For pupils in the untreated group, the reverse sequence is preferable.

  vii) Substitution effects show that for the treated group, implementing school meals until a time and replace it with a deworming program is more beneficial in terms of enhancing scores compared to the reverse.

Summary of findings (cont'd)

Treatment effects (con't)

- Enrollment, promotion and dropout:

  viii) Enrollment: exclusive negative ATE effect for the deworming program, an exclusive positive effect of the meals program, a negative overall effect and an additional positive effect.

  If we set as target the increase of enrollment, the implementation of meals program alone is preferable to deworming or the package.

  ix) Promotion and dropout: When the objective is to increase the promotion rate or reduce dropout, the package is the best option.

Summary of findings

Policy analysis: Cost-effectivemess and welfare

i) The cost-effectiveness analysis indicates that, regarding scores, deworming is far cheaper than the meals program. It also shows that introducing the meals before deworming is more cost-effective than the reverse.

ii) As for the promotion rate, the combination of the two programs is more cost-effective than the single meals program.

iii) For the dropout rate, deworming is more cost-effective compared to the canteen and the package. However, the package is more cost effective compared to the canteen only.

iv) The welfare study shows that school meals and the package contribute to enhance household's welfare.

Conclusion

Further investigation

- Multiple outcomes framework
- Multilevel models (pupils, households, villages)
- Package of more than two treatments of nutritional programme (data exist)

Limitations

- Inflation of parameters (in the Roy model).
- Difficulty (trivariate truncation): If T_j, with $j > 2$, simulation methods are required.

THANK YOU!