Sector analysis of MDG determinants

Marco V. Sánchez
UN-DESA

Inception workshop on "**Strengthening Macroeconomic and Social Policy Coherence through Integrated Macro-Micro Modelling**",
Manila, 2-3 October, 2012.

MDG determinants

- **What is needed to get all children in school and make them complete all grades?**
  - Build more school infrastructure?
  - Improve quality of other school inputs (teachers, textbook supplies)?
  - Increase access to school by improved household income and demand subsidies?
  - All of the above?
- **What is needed to reduce child mortality?**
  - Better nutrition?
  - Expansion of immunization programs?
  - Improving maternal-child health facilities?
  - Better education?
  - All of the above?
- **Are there synergies across the MDGs?**
- **What is the direct cost of interventions to achieve MDGs?**
- **Are there diminishing marginal returns to the inputs?**

Assessing MDG determinants

- No single route: country-specific determinants of MDGs
- Needs assessments and cost-effectiveness analyses
- Not just a matter of increasing public services in social sectors (i.e. more social expenditures)
  - Demand factors matter
  - Efficiency and quality of supply matters
  - Economy-wide effects

Determinants of MDG outcomes in MAMS

MDG	Service per capita or student	Consump-tion per capita	Wage incen-tives	Public infra-structure	Other MDGs
2–Primary schooling behaviour	x	x	x	x	4
4-Under-five mortality	x	x		x	7w,7s
5-Maternal mortality	x	x		x	7w,7s
7w-Water	x	x		x	
7s-Sanitation	x	x		x	

Key requirements to calibrate MDG block in MAMS

- For each MDG indicator (or schooling outcomes), two pieces of information areneeded:
  - a set of initial ***elasticities*** with respect to the determinants
    - the model recalculates the initial elasticities in consistency with the shape of the logistic functions.
  - a set of so-called ***MDG-scenario parameters***
    - provide starting points to project an outcome for an expected year, given the set of values for determinants under which—it is believed—the projected outcome could be achieved.

But also to elaborate country story on what are the factors that matter most for MDG achievement

MAMS elasticities for education

$$ZEDU_{b,c,t} = \alpha educe_{b,c} \cdot \left(EDUQUAL_{c,t} \right)^{\varphi edu_{b,c,edu-qual}} \cdot \left(\frac{WF_{f-labs,t}}{WF_{f-labn,t}} \right)^{\varphi edu_{b,c,w-prem}} \cdot \left(\frac{WF_{f-labt,t}}{WF_{f-labs,t}} \right)^{\varphi edu_{b,c,w-prem}} \cdot MDGVAL_{mdg4,t}^{\varphi edu_{b,c,mdg4}} \cdot \prod_{f \in FCAPGOVINF} \left(\sum_{i \in INS} QFINS_{i,f,t} \right)^{\varphi edu_{b,c,f}} \cdot QHPC_t^{\varphi edu_{b,c,qhpc}}$$

$$\left[\begin{array}{l} intermediate\ variable\ for\ student \\ share\ with\ behavior\ b\ in\ cycle\ c \end{array} \right]$$

$$= \left[\begin{array}{l} exogenous \\ trend\ value \end{array} \right] \cdot \left[\begin{array}{l} influence\ of: education\ quality\ (service\ per\ student); \\ wage\ premia\ (for\ c \leq secondary\ and\ c \geq tertiary, resp.); \\ student\ health\ (proxied\ by\ MDG4);\ level\ of\ infra- \\ structure;\ and\ per\text{-}capita\ household\ consumption \end{array} \right]$$

MAMS elasticities for other MDGs

$$ZMDG_{mdg,t} = \alpha mce_{mdg} \cdot \left(\prod_{cmdg \in CMDG} \left(\sum_{\substack{c \in C \\ |(cmdg,c) \in MCM}} \frac{QQ_{c,t}}{poptot_t} \right)^{\varphi m_{mdg,cmdg}} \right)$$

$$\cdot \prod_{f \in FCAPGOVINF} \left(\sum_{i \in INS} QFINS_{i,f,t} \right)^{\varphi m_{mdg,f}}$$

$$\cdot \left(\prod_{mdg' \in MDGSTD} MDGVAL_{mdg',t}^{\varphi m_{mdg,mdg'}} \right) \cdot QHPC_t^{\varphi m_{mdg,"hhdconspc"}}$$

$$\begin{bmatrix} intermediate\ variable \\ for\ MDGs\ 4\ and\ 5 \end{bmatrix} = \begin{bmatrix} exogenous \\ parameter \end{bmatrix} \cdot \begin{bmatrix} influence\ of: real\ value\ for\ services\ per\ capita; \\ level\ of\ infrastructure;\ water\ and\ sanitation\ MDGs; \\ household\ consumption\ per\ capita \end{bmatrix}$$

How to obtain "MDG elasticities"?

1. Most desirable practice: own estimates, data permitting
2. If 1 is not possible, use values already estimated and available
   - Must be up to good econometric standard
   - With the right disaggregation
   - Rarely an option for MAMS
3. Less desirable (but some times the only available) practices
   - borrow values from existing studies
   - use 'educated' guesses (really least desirable!)

Whatever the route you take:

- do sensitivity analysis of simulation results to changes in elasticity values
- talk to "sectoral experts"
- making sure MDG trends in MAMS are plausible

How to estimate?

Econometric specifications

- Probability model of different forms (logit, probit, Multi-nomial logit)
  - Probability of attending school given socio-economic conditions of household, individual characteristics (gender, ethnicity, nutrition) and quality of supply inputs
  - MNL if there is a choice between, say, private and public education
- Proportions model: estimate rate of enrolment or graduation rate directly
  - Logit quasi-maximum likelihood methodology (OLS not appropriate)
  - Estimate proportions, e g. across provinces, municipalities or districts. You may lose some variability

How to estimate?

The logit model

$$\Pr\left(Y=1 \middle| x_i\right) = F\left(x_i \beta\right)$$

i **:** independent variable for *x*

Y **:** dependent variable (i.e. MDG indicator), taking a value of 1 or 0.

F( ) **:** standard logistic function

x_i **:** contains vectors of relevant socio-economic factors thought to affect the *Y* variable.

β **:** estimated coefficient in logit model

From estimated coefficients to elasticities

The logit model

β **:** estimated coefficient in logit model

• Estimated coefficients do not have a direct economic interpretation.

• Economists use marginal effects and elasticities.

• Marginal effects of independent variables calculated using beta: the probability that determinant X affects Y is #

• Elasticities needed to calibrate MAMS!

The logit model

An elasticity gives the % change in the probability of a success when the explanatory variable has changed by 1%. For the *i* explanatory variable, this is obtained using partial derivatives as:

$$\frac{\partial \Pr\left(Y=1\middle|x_i\right)}{\partial x_i}\cdot\frac{x_i}{\Pr\left(Y=1\middle|x_i\right)} \qquad \varepsilon = \frac{\partial Y}{\partial x_i}\cdot\frac{x_i}{Y}$$

- The elasticities vary for every observation *j*: logit models usually work for individuals/individual households; i.e. ε_j.
- Summary measure needed: i.e., the sample means of the explanatory variables. In the last equation, if *j* represents *n* individuals or households, the elasticity is:

$$\varepsilon = \frac{\sum_{j=n}\varepsilon_j}{n}$$

Elasticities for the determinants of MDG 2 in MAMS

MDG	Service per capita or student	Consump-tion per capita	Wage incen-tives	Public infra-structure	Other MDGs
2–Primary schooling (outcomes)	**x**	**x**	**x**	**x**	**4**
4-Under-five mortality	x	x		x	7w,7s
5-Maternal mortality	x	x		x	7w,7s
7w-Water	x	x		x	
7s-Sanitation	x	x		x	

Education-related elasticities

- What dependent variable(s)?
  - Probability of entering primary school (*neting1*)
  - Probability of passing a given grade by cycle (*prom*)
  - Probability of graduating from previous cycle and continuing to the next (*grdcont*)
- What independent variables?
  - MAMS determinants + other control variables
  - Demand factors: household income, education level of parents, and so on.
  - Supply factors: geographical accessibility to school, quality of school inputs (qualified teachers, test scores, pupil-teacher ratio, etc.)
- Data requirements for estimation:
  - Household survey data;
  - Data on government spending in education sector (by governorate or other administrative unit).

Ecuador - Logit model	Marginal effect	Elasticity	p-value
	Prob of primary enrolment (grdentry)		
Consumption per capita	0.00000046	0.126	0.001
MDG4	-0.00004750	-0.035	0.166
Education quality (services)	0.00077250	0.111	0.143
Public Infrastructure	0.18224220	0.162	0.023
Wage premium (W_2 / W_1)	0.03375350	0.059	0.193
	Prob of graduating primary (grdp)		
Consumption per capita	0.00000012	0.030	0.005
MDG4	-0.00001930	-0.013	0.169
Education quality (services)	0.00036280	0.050	0.052
Wage premium (W_2 / W_1)	0.02430020	0.041	0.027
	Prob of continuing to secondary (grdcons)		
Consumption per capita	0.00000027	0.087	0.000
MDG4	-0.00002670	-0.019	0.157
Public Infrastructure	0.10860630	0.086	0.048
Wage premium (W_2 / W_1)	0.02436420	0.034	0.119
	Prob of continuing to tertiary (grdcont)		
Consumption per capita	0.00000017	0.097	0.148
Public Infrastructure	0.74773540	0.821	0.016
Wage premium (W_3 / W_2)	0.06347780	0.203	0.199
	Prob of graduating secundary and tertiary		
MDG4	-0.00003100	-0.025	0.144
Education quality (services)	0.01011030	0.253	0.003
Public Infrastructure	0.09554830	0.080	0.255
Wage premium (W_3 / W_2)	0.02661770	0.046	0.136

Other determinants in model specification:

- Education input indicators (pupils/class room; quality teachers; degree of school autonomy)
- Parents education
- Other control variables (urban/rural, residence, ethnicity, and others)

An econometric specification and empirical results for Yemen

- Determinants of enrolment behaviour estimated for Yemen using the Household Budget Survey for 2005/2006.
- Choice of the estimable specification is based on literature review and it also follows the specification of MAMS for student behaviour.
- The survey dataset allowed to estimate student behaviour by cycle only for entry (or enrolling for the first time) and enrolment rates since the survey dataset lacks detail on students passing, failing or repeating.
- Even so, the empirical results can provide a good reference point to assign initial elasticity values to MAMS.
- Dependent variable takes a value of 1 if the individual–of the relevant age cohort for the cycle—attended school at the time when the survey was conducted, or 0 otherwise.

$$y = \alpha_1\ Area + \alpha_2\ Sex + \alpha_3\ Head\_edu + \alpha_4\ Spouse\_edu + \alpha_5\ Health$$
$$+\alpha_6\ Inc\_pc + \alpha_7\ Inf + \alpha_7\ Edu\_qual + \alpha_8\ Wage\_prem$$

Logistic regression results for entry and attendance in Yemen's basic education

	Entry			Attendance		
	Parameter estimates	Marginal effects	Elasticities	Parameter estimates	Marginal effects	Elasticities
Sex	0.503 (4.04***)	0.120	0.105	1.300 (20.3***)	0.217	0.140
Area	-0.597 (-3.91***)	-0.146	-0.057	-0.046 (-0.60)		
Head_edu	0.455 (3.29***)	0.109	0.086	0.551 (7.86***)	0.089	0.052
Spouse_edu	0.454 (2.33**)	0.105	0.030	0.582 (5.13***)	0.084	0.018
Edu_ qual	0.251 -1.26			0.987 (9.14***)	0.162	0.148
Inc_pc	1.21 (3.35***)	0.291	0.485	1.51 (8.05***)	0.247	0.312
Inf	1.386 (1.26***)	0.333	0.726	1.35 (19.03***)	0.223	0.394
Health	-0.75 (-1.77*)	-0.187	-0.008	-0.978 (-4.68***)	-0.201	-0.005
Wage_prem	-0.089 (-0.36)			-0.353 (-2.90***)		

Elasticities for the determinants of MDGs 4 & 5 in MAMS

MDG	Service per capita or student	Consump-tion per capita	Wage incen-tives	Public infra-structure	Other MDGs
2–Primary schooling behavour	x	x	x	x	4
4-Under-five mortality	**x**	**x**		**x**	**7w,7s**
5-Maternal mortality	**x**	**x**		**x**	**7w,7s**
7w-Water	**x**	**x**		**x**	
7s-Sanitation	**x**	**x**		**x**	

Estimating mortality-related elasticities for MAMS

- **What dependent variable per equation?**
  - Under-five mortality (or infant mortality): binary variable, indicating if a child died under the age of five (or one)
  - Maternal mortality: binary variable, indicating if a woman died of child-birth related reason; not discussed here, similar approach…

Estimating under-five mortality-related elasticities for MAMS

- **What independent variables?**
  - Child characteristics (personal and biological);
  - Maternal (behavioural) characteristics;
  - Socio-economic household and community characteristics:
    - including MAMS determinants:
      - Access to safe drinking water (MDG 7w);
      - Access to improved sanitation facilities (MDG 7s);
      - Per capita household consumption (or proxy);
      - Per capita expenditure on health services;
      - Other public infrastructure.
- **Data requirements for estimation:**
  - Household survey data (e.g. DHS), covering household, maternal and child characteristics;
  - Data on government spending in the health sector (by governorate or other administrative unit).

Estimating under-five mortality-related elasticities for MAMS – cont.

- **Some possible data problems:**
  - Lack of information on per capita household consumption? → Use proxy: DHS "wealth index"
  - Availability of data on government health spending in base year?
  - MAMS only captures the same-year effects of gov't spending, no room for lagged effects! If spending is targeted at underperforming areas → reverse causation bias!
  - Measure for "other public infrastructure"? If data on public transport infrastructure not available, consider using "access to electricity". Problematic if all household have access – use alternative proxy, such as "access to water all day"

Commonly used estimation techniques

- **Logit/probit model:**
  - Discrete binary dependent variable
  - Continuous or discrete (dummy) determinants
  - Assumption about underlying probability distribution: logistic (logit) or normal (probit) $\rightarrow$ use specification test to decide!
- **Cox proportional hazard model (survival model):**
  - No assumptions about functional form of underlying distribution (hazard function), only "proportional hazards" assumption;
  - Better use of information than logit/probit, isolating effect of age (in months) on mortality from effect of exogenous determinants;
  - Handling of censored data, using information from "cut-off" cases (children who are alive and less than five/one years old at time of survey);
  - More detailed information required about age at death, and computationally more complex…
  - Proportional hazards assumption may prove restrictive!

The logit model

$$\text{Prob}(Mort = 1 \mid \mathbf{x}) = F(\mathbf{x}'\boldsymbol{\beta})$$

Mort **:** MDG indicator

- for under-five/infant mortality, taking a value of 1 if a child died at less than five/one years of age and 0 otherwise
- for maternal mortality, taking a value of 1 if a woman died for child-birth related reasons and 0 otherwise

x **:** vector of relevant socio-economic factors thought to affect child or maternal mortality

F( ) **:** standard logistic function

β **:** vector of coefficients

Cox-Proportional-Hazard (CPH) survival model

$$H_j(t) = e^{\sum_{ij} \beta_j x_i} H_0(t)$$

- $H_i(t)$: risk of child '*j*' to die in period (*t*) before reaching 5 years of age;
- $H_0(t)$: risk of child of reference group to die in period (*t*) before reaching 5 years of age;
- x_i: determinants of child mortality.

Example: Determinants of infant mortality in Jordan (1)

- Based on 2007 DHS for Jordan
- Estimation for infant mortality – more relevant in Jordan (under-five mortality yields similar results)
- Cox-Hazard model can't be used to assess gender bias (violation of proportional hazards assumption)
- Specification test $\rightarrow$ use logit model!!

Example: Determinants of infant mortality in Jordan (2)

Jordan results

Elasticities for the determinants of MDG7, targets w & s in MAMS

MDG	Service per capita or student	Consump-tion per capita	Wage incen-tives	Public infra-structure	Other MDGs
2–Primary schooling behaviour	x	x	x	x	4
4-Under-five mortality	x	x		x	7w,7s
5-Maternal mortality	x	x		x	7w,7s
7w-Water	**x**	**x**		**x**	
7s-Sanitation	**x**	**x**		**x**	

An econometric specification and empirical results for Yemen

- Demographic and Health Survey (DHS), conducted for 2003 by the Central Statistical Organization of the Ministry of Planning and International Cooperation of the Republic of Yemen.
- Two specifications were used to conduct the econometric estimations for, respectively, access to improved drinking water and access to improved sanitation facilities.
- Individuals surveyed who claimed they had piped or cooperative supply, artisan or regular well, or bottled water, were considered to have access to improved drinking water. Dummy = 1.
- Improved sanitation facilities are considered to be a flush toilet connected or not connected to sewage, a pit, or a toilet with tank. Dummy =1.
- Estimated specification for the dummy variables (y_i) :

$$\Pr ob(y=1) = \frac{\exp^{\alpha*wealth+\beta*area+\gamma*spending\_pc+\theta*eletricity}}{1+\exp^{\alpha*wealth+\beta*area+\gamma*spending\_pc+\theta*eletricity}}$$

Logistic regression results for water and sanitation in Yemen

	Water equation			Sanitation equation		
	Parameter estimates	Marginal effects	Elasticities	Parameter estimates	Marginal effects	Elasticities
wealth	1.265 (22.94)	0.246	0.077	5.997 (85.68)	0.961	0.554
area	0.137 (5.46)	0.026	0.009	1.141 (43.66)	0.203	0.129
spending_pc	0.056 (2.35*)	0.011	0.015	0.272 (9.62)	0.043	0.104
electricity	0.604 (8.34)	0.117	0.071	1.135 (13.07)	0.182	0.194

The following notes apply to this table: (i) z-statistics are presented in brackets; (ii) the statistical significance is at the 1% in all cases but those where an asterisk has been added; (iii) the marginal effects are defined as $\Delta y/\Delta x$, where Δ denotes change, y is the value of the dependent variable, and x represents the value of the determinant (s); and, the elasticity is computed as follows: $(\Delta y/y)/(\Delta x/x)$.

When estimating, keep in mind:

- Data need to be carefully explored!
- Estimates may be sensitive to model specification:
  - are we using the correct variables and are these well represented by the data?
  - are we using the correct model specification?
  - use of proxy variables or dummy variables to control for time and space.
- Possible endogeneity problems (e.g. distribution of public education spending may be determined by enrolment rates; targeting of health spending at lagging areas);
- Deal with multicollinearity (e.g. per capita consumption and infant mortality may be correlated);
- Causality not being really tested! Only correlations!
- Carefully interpret results and link back to the estimation stage, and check whether elasticity is plausible or not, before settling on final results!

… and some more

- Estimated elasticities can not necessarily be applied one-to-one for calibrating MAMS:
  - estimated models tend to be better specified;
  - independent variables used in estimation differ from those used in MAMS;
  - ➢ Need to complement with knowledge of sector experts!
- Estimated elasticities should be a *starting point* for running MAMS. Debugging will most likely be needed to generate reasonable results!!
  - Validate the values by examining the trends of MDG indicators in baseline scenario!

We have an Unhappy Marriage

- … between Mr. ECONS (econometrics) and Ms. MAMS (CGE model)
- Prenuptial agreement: all is on MAMS' (CGE model's terms)
  - Drop stochastic elements of the estimation (i.e. all becomes deterministic)
  - Forget about other determinants that are not in CGE model (though estimation controls for them)
  - Don't worry about explanatory power
- If you can accept these prenuptial terms, it might be a good marriage after all (happy it will never be…)