Plenary Session IV:

Suggestions for estimating MDG-related elasticities

Cornelia Kaldewei   Marco V Sánchez

UN-DESA

Presentation for the Third Training Workshop of the Project "Assessing Development Strategies to Achieve the MDGs in Asia", Jakarta, March 30 – April 2, 2010

Try to estimate as much as possible

1. Most desirable practice: estimate your elasticities as much as data allow it

2. If 1 is not possible, use values already estimated and available
   - These must be up to good econometric standard
   - Make sure disaggregation is useful
   - Rarely an option for MDG-related elasticities in MAMS

3. Less desirable (but some times the only available) practices
   - borrow elasticity values from existing studies
   - use 'educated' guesses (really least desirable!)

- Whatever the route taken, do an effort to validate through sensitivity analysis of simulation results to changes in elasticity values
  - Talking to "sectoral experts" has proven useful to define plausible values.

MDG elasticities: found in <app>-data-mdg.xls

- ***mdgeduelas(ac,acp,acpp)***: elasticity of indicator (1st two indices) with respect to determinant (3rd index)
  - Rarely found already estimated.
  - Need to be estimated
    - Logit and probit models
    - Survival models for mortality
  - Purpose: complete requirements of the dataset but also <u>elaborate country story on MDG determinants</u>

		c-hlthg	c-wtsn	edu-qual	f-capoinf	qhpc	mdg4	mdg7a	mdg7b	wage-prem
mdg1	dummy					-1.000				
mdg4	dummy	-0.485			-0.048	-0.048		-0.097	-0.097	
mdg5	dummy	-0.864			-0.086	-0.086		-0.086	-0.086	
mdg7a	dummy		0.287		0.029	0.057				
mdg7b	dummy		0.644		0.129	0.064				
g1entry	c-edup1			1.100	0.110	0.110	-0.110			0.110
pass	c-edup1			0.867	0.087	0.087	-0.087			0.087
pass	c-edup2			0.119	0.012	0.012	-0.012			0.012
pass	c-edus			0.171	0.017	0.017	-0.017			0.017
pass	c-edut			0.137	0.014	0.014	-0.014			0.014
grdcont	c-edup2			1.049	0.105	0.105	-0.105			0.105
grdcont	c-edus			0.203	0.020	0.020	-0.020			0.020
grdcont	c-edut			1.231	0.123	0.123	-0.123			0.123

MDG related indicator

mdg4	under-five mortality rate
mdg5	maternal mortality rate
mdg7a	proportion of population using an improved drinking water source
mdg7b	proportion of population using an improved sanitation facility
mdg2	primary completion rate
g1entry	(share of) 1st-year primary cohort that enter first primary cycle
pass	(share of) student passing in each grade (excluding those students graduating from cycle)
grdcont	(share of) graduates from previous cycle who continue to next cycle

MDG related determinant

c-hlthg	real government spending per capita - health
c-wtsn	real government spending per capita - water & sanitation
edu-qual	real government spending per student - education (proxy for quality of education)
f-capoinf	level of public infrastructure (public roads, bridges, airports, etc.)
qhpc	real household consumption per capita
mdg4	under-five mortality rate (proxy for student health)
mdg7a	proportion of population using an improved drinking water source
mdg7b	proportion of population using an improved sanitation facility
wage-prem	wage premium:
	the wage of those with complete primary education relative to those with incomplete or none education
	the wage of those with complete secondary education relative to those with complete primary education
	the wage of those with complete tertiary education relative to those with complete secondary education

How to estimate?

The logit model

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

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.

   : estimated coefficient in logit model

From estimated coefficients to elasticities

The logit model

: estimated coefficient in logit model

• For the logit model, the estimated coefficients do not have a direct economic interpretation.

• Measures that are familiar to economists are 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 percentage change in the probability of a success in response to a one percentage change in the explanatory variable. For the *i* explanatory variable this is obtained using partial derivatives as:

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

- The elasticities vary for every observation: logit models usually work for individuals or 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:

$$\varepsilon = \frac{\sum\limits_{j=n} \varepsilon_j}{n}$$

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	**X**	**X**	**X**	**X**	**4**
4-Under-five mortality	**X**	**X**		**X**	**7a,7b**
5-Maternal mortality	**X**	**X**		**X**	**7a,7b**
7a-Water	X	X		X	
7b-Sanitation	X	X		X	

Estimating education-related elasticities

- What dependent variable(s)?

  - Probability of entering primary school (*grd1entry*)
  - Probability of passing a given grade by cycle (*grd*)
  - Probability of graduating from previous cycle and continuing to the next (*grcont*)

- What independent variables?
  - MAMS determinants + other control variables
  - Demand factors: household income, education level of parents, and so on.
  - Supply factors: physical 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).

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	Consumption per capita	Wage incentives	Public infrastructure	Other MDGs
2–Primary schooling	X	X	X	X	4
4-Under-five mortality	X	X		X	7a,7b
5-Maternal mortality	X	X		X	7a,7b
7a-Water	X	X		X	
7b-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 7a);
      - Access to improved sanitation facilities (MDG 7b);
      - Per capita household consumption (or proxy);
      - Per capita expenditure on public health services;
      - Other public infrastructure.

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

- **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 – ctd.

- **Some possible data problems:**
  - Lack of information on per capita household consumption? $\rightarrow$ 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 $\rightarrow$ reverse causation bias!
  - Measure for "other public infrastructure"? If data on public transport infrastructure not available, consider using "access to electricity". Problematic if all hh 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

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

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)

- Refer to documentation in "tool kit" (technical note, list of variables, data base, Stata do-files)
- 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)

- Standard Stata estimation and post-estimation commands:

```
logit under1 var_1 … var_n if exit1probit == 0
```

`mfx`   **(marginal effects)**

`mfx, eyex`   **(elasticities)**

Note: marginal effects and elasticities are evaluated at the means; for dummy variables at change from 0 to 1.

Example: Determinants of infant mortality in Jordan (3)

Jordan results

Elasticities for the determinants of MDG7, targets a & b in MAMS

MDG	Service per capita or student	Consumption per capita	Wage incentives	Public infra-structure	Other MDGs
2–Primary schooling	X	X	X	X	4
4-Under-five mortality	X	X		X	7a,7b
5-Maternal mortality	X	X		X	7a,7b
7a-Water	X	X		X	
7b-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: $(\Delta y / y)/(\Delta x / x)$.

12. 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.
- Deal with 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);
- 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!

Suggested references

- Kaldewei, Cornelia. 2010. "Determinants of Infant and Under-Five Mortality – The Case of Jordan" Development Policy and Analysis Division, Department of Economic and Social Affairs, New York (unprocessed, March version).

- Lofgren Hans. 2008. "MAMS – A Guide for Users". December 6 (unprocessed).

- Sabrana, Giacomo. 2009. "Technical note on the determinants of water and sanitation in Yemen", Development Policy and Analysis Division, Department of Economic and Social Affairs, New York (unprocessed, June version).

- Sánchez, Marco V. and Giacomo Sbrana. 2009. "Determinants of education attainment and development goals in Yemen", Development Policy and Analysis Division, Department of Economic and Social Affairs, New York (unprocessed, May version).