# Going beyond functionings to capabilities: an econometric model to explain and estimate capabilities

by

Jaya Krishnakumar

University of Geneva

August 2005

International Conference on

The Many Dimensions of Poverty

International Poverty Centre

Brasilia

29-31 August 2005

An attempt

- to understand the complex mechanisms involved in capability enhancement
- by providing a theoretical structure for their representation, explanation and measurement
- leading to an econometric model that can be estimated using real data
- yielding estimators that reflect capabilities derived from functionings

# Outline of the presentation

Introduction

Simultaneity

Measurement

Theoretical Framework

Econometric Model

Empirical Application

Conclusions

# Introduction

The basic purpose of development is to enlarge people's choices so that they can lead the life they want to (Sen (1985, 1999)).

Development or welfare is a multidimensional concept enveloping diverse social, economic, cultural and political dimensions

Capabilities are the *choices* that one has

Functionings are the actual outcomes or the levels of achievement attained in the various dimensions

Capabilities are unobservable

Functionings are observable/measurable

Therefore an appealing framework is:

A model which assumes that the capabilities are *latent* variables *manifesting* themselves through a set of observed indicators.

Examples of such models used in our context: Principal components, factor analysis and MIMIC (multiple indicators and multiple causes)

Principal components:

The latent variables are estimated as linear combinations of the observed indicators chosen in such a way as to reproduce the original data as closely as possible.

But there is no underlying theoretical model.

Factor analysis:

Offers a theoretical explanation: the observed values are postulated to be (linear) functions of a certain number (fewer) of unobserved latent variables (called factors). Thus it provides a framework for going beyond functionings to reach the capabilities represented by the latent factors.

But this model does not *explain* the latent variables (or the capabilities in our context) in that it does not say what causes these capabilities to change.

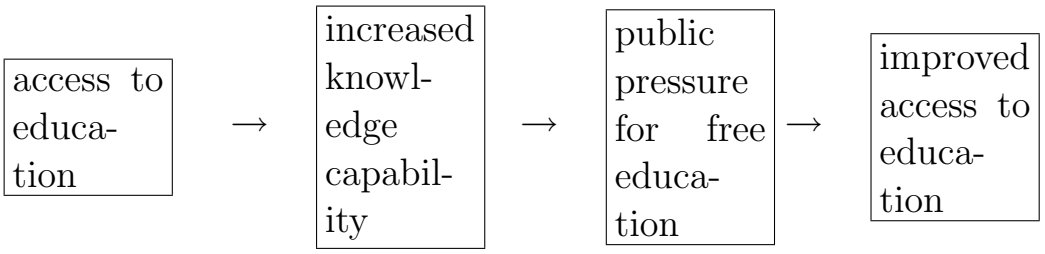
MIMIC model:

Here it is not only believed that the observed variables are manifestations of an underlying unobserved latent concept but also that there are other exogenous variables that "cause" and influence the latent factor(s).

This model is a step ahead in the right direction but...

We believe: Not only do these "causes" influence capabilities or human development but they are also influenced by it.

A simple example :




Other relevant models for the above situation:

LISREL with ordinal variables

MIMIC with exogenous variables

But *the capability approach* needs more...

## The Simultaneous Nature of Capabilities

Remember capabilities are the choices that one faces in life and functionings are the outcomes.

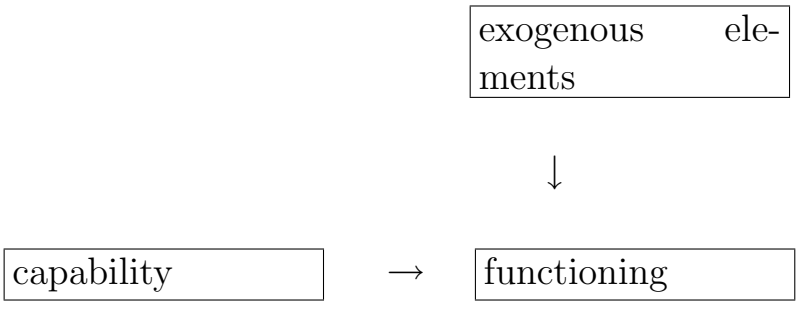
So given a capability level it is possible to have more than one achievement levels.

Take education for instance.

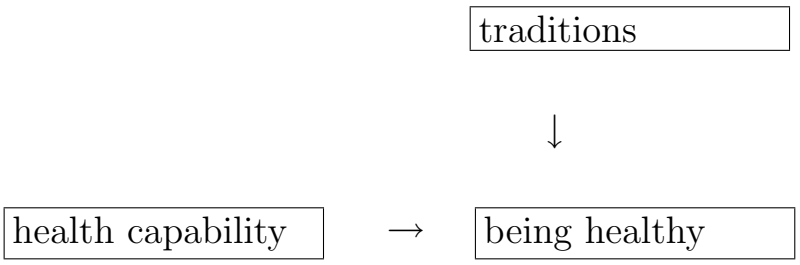
From capability to more than one achievements

| | | |
|---|---|---|
| knowledge capability | → | being educated |
| knowledge capability | → | not being educated |

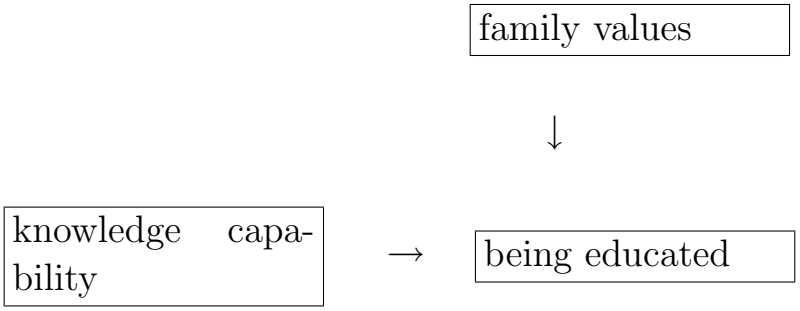
Why: because of exogenous elements



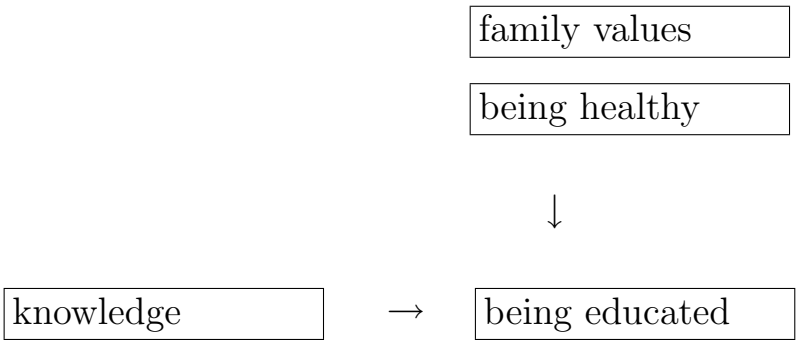
For instance in the case of health




Similarly in the case of education

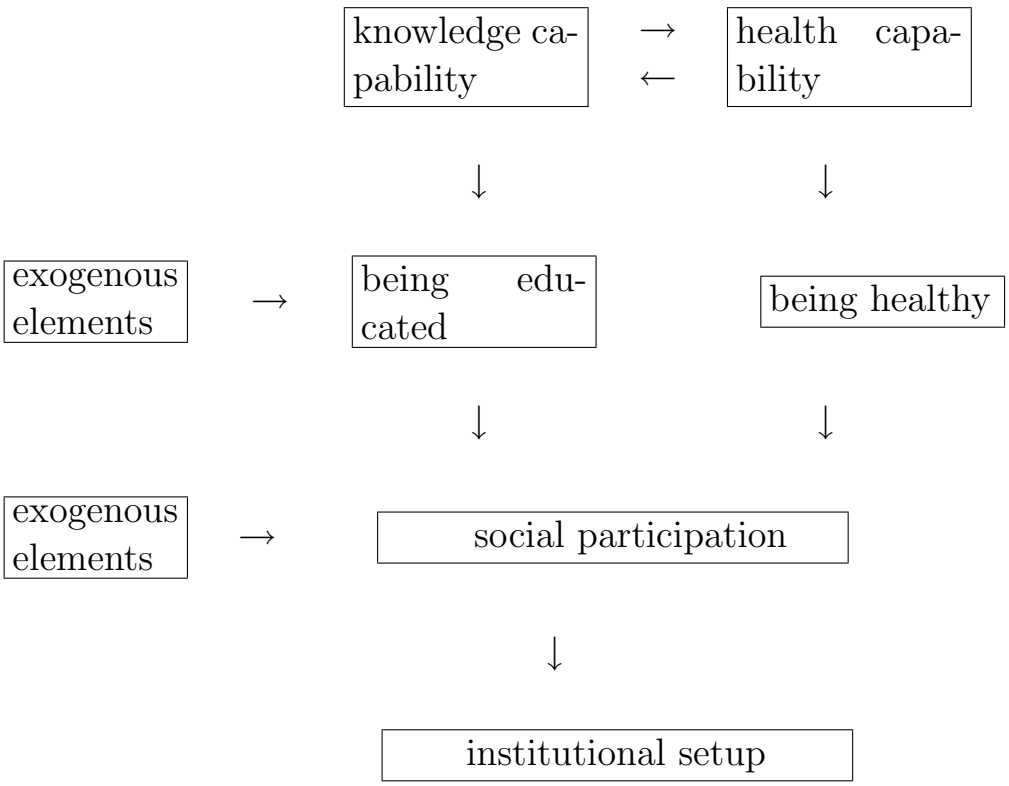



But we could also add 'being healthy' as an influencing factor for education or vice versa: interdependence



Now,

Education and health encourage social participation




Increased social participation in turn affects the institutional setup: so some of these 'exogenous' elements not so 'exogenous' after all!

Therefore:

Capabilities are interdependent

Not only do capabilities interact among themselves but also with some external elements representing the socio-political setup.

Thus there are feedback effects; some external factors are in fact potentially endogenous.

For other factors like individual characteristics, traditions, culture, the causal link only operates in one direction (making them purely exogenous).

## The Measurement Issue

Capabilities by definition cannot be directly measured.

Hence they need to be specified as latent unobservable variables in our model.

What can be measured however are the functionings namely the achievements in each dimension both at the individual (household) and at the national levels. These achievements are generally identified by proper indicators reflecting the performance in the associated dimension.

One normally has a vector of functionings rather than a scalar indicator corresponding to each domain.

For example in the case of health, at the national level, one can think of classic indicators such as life expectancy, infant mortality, child mortality and in the case of education literacy rate, enrolment ratio etc.

# Types of indicators available in practice:

Quantitative or Continuous

Most common like the above-mentioned life expectancy, per capita number of doctors etc.

Qualitative

*Binary* or *dichotomous*: two outcomes

For instance the existence of the right to vote or not, existence of safe water access or not, existence of adequate sanitation facilities or not, also subjective assessments like whether a person considers herself to be poor or not.

*Polychotomous*: more than two outcomes

Ordinal or sequential: e.g. different levels of education - no formal education, primary, secondary, college...

With no order for example religion - Hindu, Muslim, Buddhist, Christian etc.

*Truncated or censored*

Truncated when not observed for a particular range of values, censored when observed only if greater than a threshold value.

## The General Theoretical Framework

(i) Capabilities are *latent, unobservable* and interdependent, and are *endogenous* in our structural model.

(ii) Capabilities are influenced by a set of social, political and institutional factors some of which may in turn be influenced by them. (In addition to capabilities there are also some *observed endogenous* variables in our model.)

(iii) Capabilities are also influenced by a set of observable external/*exogenous* causes (such as traditions, cultural elements, natural environmental factors and some social, political, institutional ones which are not part of (ii)).

(iv) Achievements/functionings are measurable and are linked to the underlying capabilities (the set of relationships linking the two is the so-called measurement model or the qualitative response model).

(v) The relationships between the latent capabilities and the observed functionings are also affected by *exogenous* elements (for instance individual characteristics).

# Some notations

$y^*$ a vector of latent unobserved capabilities say $(m \times 1)$

$y$ a vector of observed indicators representing the functionings associated with the capability vector say $(p \times 1)$

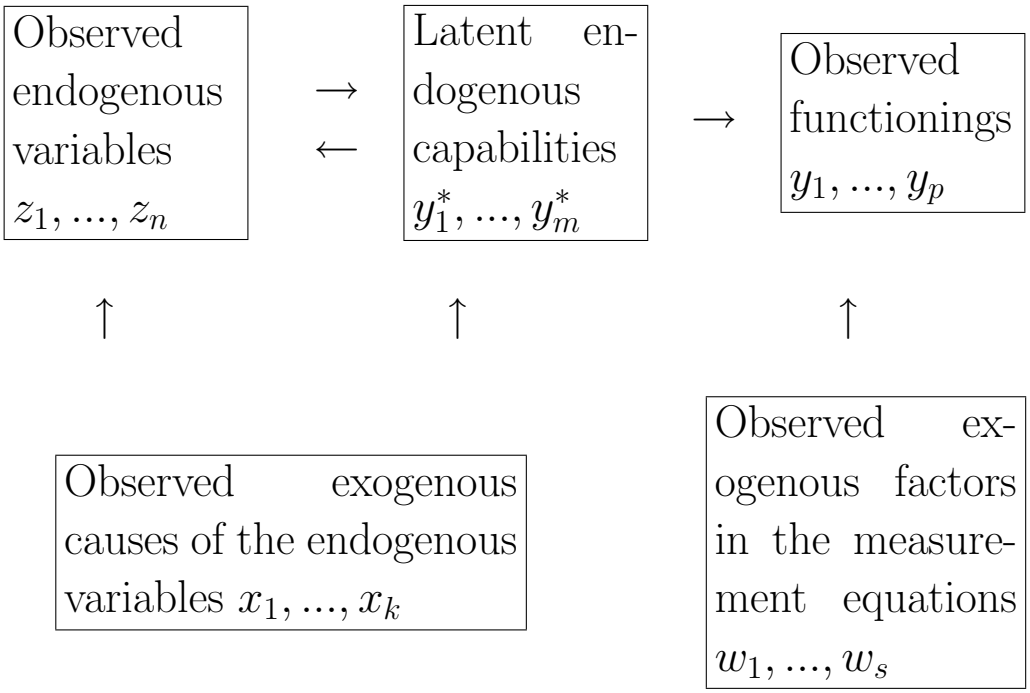
as discussed earlier, some these $y$'s could be continuous, some qualitative or discrete

$z$ a vector of observed variables that influence the capabilities but are also influenced by them say $(n \times 1)$

$x$ a vector of exogenous "causes" of $y^*$ and $z$ say $(k \times 1)$

$w$ a vector of exogenous factors entering the measurement equations i.e. the relationships between observed indicators $y$ and latent variables $y^*$ say $(s \times 1)$

# Our framework:



## The Econometric Model

$$Ay^* + Bz + Cx + u = 0 \tag{1}$$

$$g(y) = h(y^*, w) + v \tag{2}$$

Equations (1) represent the structural simultaneous equation model (SEM) which jointly explains $(y^*, z)$ in terms of $x$, with $A$, $B$, $C$ being the corresponding coefficient matrices of appropriate dimensions.

Equations (2) form the measurement model or the qualitative response model (QRM) where it is specified how the latent variables are related to the observed responses through functions $g(\cdot)$ and $h(\cdot)$. Note the presence of exogenous variables in both the models.

Estimation:

By two-stage methods via the reduced form (see example)

# Some special cases:

Case 1

If $y$ is continuous, $g(\cdot), h(\cdot)$ linear and there is no $w$ we get the standard LISREL model (cf. Joreskog (1973)) (with observed rather than latent exogenous, refer to an earlier remark in this respect).

Case 2

With ordinal $y$ and no $w$ we have LISREL with ordinal variables (cf. Joreskog (2002), Muthen (1983, 1984)). The latter author has two types of measurement equations: 'inner' measurement equations and 'outer' measurement equations as he allows for *latent* response variables and *observed* response variables.

Case 3

If $y^*$ scalar, $A = 1$, no $z$, no $w$, $y$ continuous we have the MIMIC model (cf. Joreskog and Goldberger (1975)).

Case 4

Same as Case 3 with $y^*$ a vector, $A = I$, we have the extended or generalised MIMIC.

Case 5

Same as Case 4 with $w$ and $z$, we have the MIMIC with covariates (cf. Moustaki (2003)).

Case 6

If $y^*$ is observed (no measurement equation) then we have the classical SEM (cf. e.g. Theil (1979), Hausman(1983)).

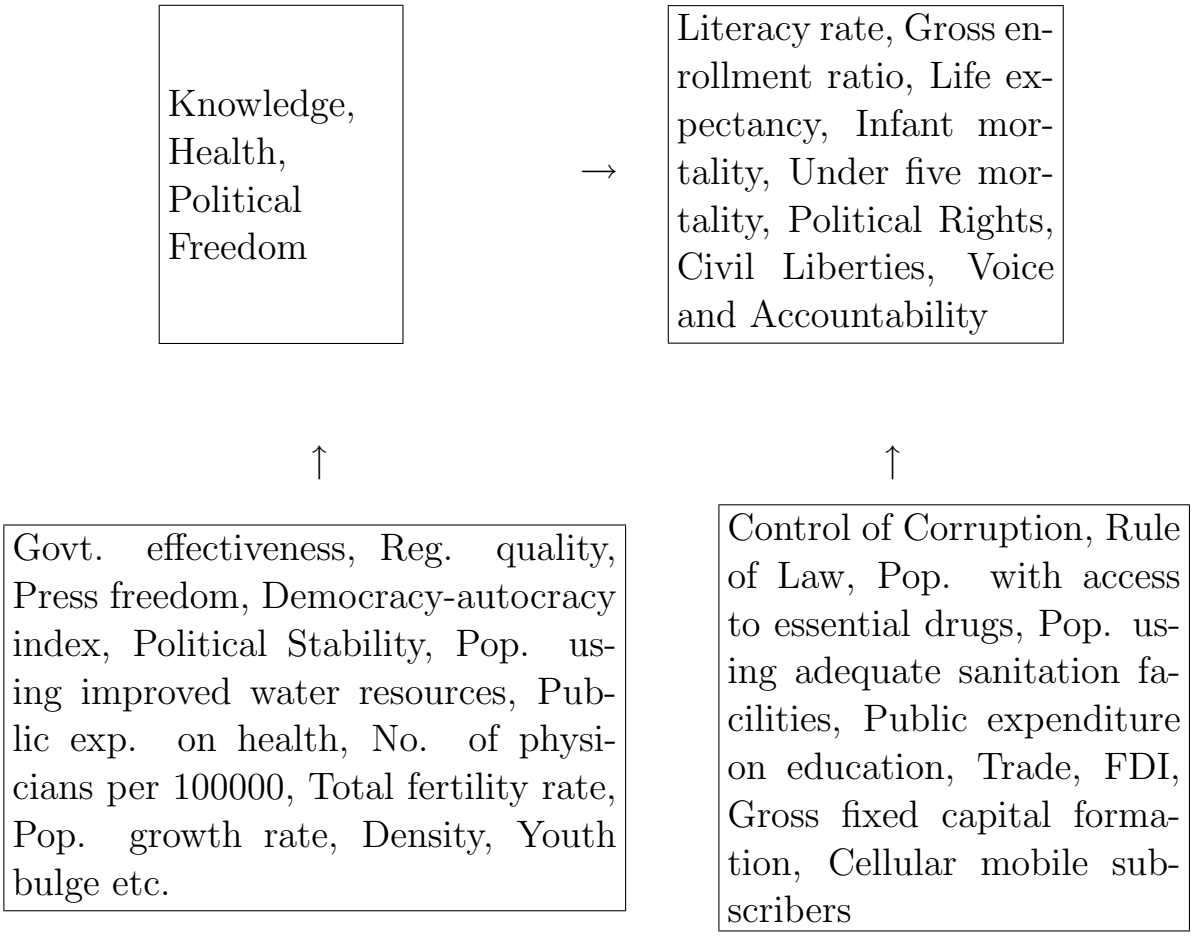
Case 7

If $y^*$ is observed and $A = I$, then we have the SUR model (cf. Zellner (1962)).

Case 8

When $y^*$ is scalar (no $z$) and $y$ is either discrete or limited dependent we have the classical qualitative dependent variable model (see Amemiya (1985)).

# Empirical Model



## Data Sources

| | |
|---|---|
| UNDP: | Human Development Data |
| World Bank Group: | World Development Indicators |
| World Bank Group: | Worldwide Governance Research Indicators |
| CIFP: | Risk Assessment Indicators |

## List of Variables

| The latent endogenous variables | |
|---|---|
| $y_1^*$: | Knowledge |
| $y_2^*$: | Health |
| $y_3^*$: | Political Freedom |
| **The achievement indicators** | |
| $y_1$: | Political Rights |
| $y_2$: | Civil Liberties |
| $y_3$: | Voice and Accountability |
| $y_5$: | Life expectancy at birth (years) |
| $y_6$: | Adult literacy rate (% age 15 and above) |
| $y_7$: | Combined primary, secondary & tertiary gross enrolment ratio (%) |
| $y_8$: | Infant mortality rate (per 1,000 live births) |
| $y_9$: | Under-five mortality rate (per 1,000 live births) |

## List of Variables: contd.

| Possible exogenous variables (observed) | |
|---|---|
| For the structural part | |
| $x_1$: | Government Effectiveness |
| $x_2$: | Regulatory Quality |
| $x_3$: | Population using improved water sources (%) |
| $x_4$: | Cellular mobile subscribers (per 1.000 people) |
| $x_5$: | Public expenditure on health (% of GDP) |
| $x_6$: | Total debt service (% of GDP) |
| $x_7$: | Density (persons per sq.km.) |
| $x_8$: | Political Stability |
| $x_9$: | Population Growth Rate (Annual %) |
| $x_{10}$: | Urban Population Growth Rate (Annual %) |
| $x_{11}$: | Youth Bulge (Pop. Aged 0-14 as a % of Total) |
| $x_{12}$: | Physicians (per 100,000 people) |
| $x_{13}$: | Press Freedom |
| $x_{14}$: | Democracy - Autocracy Index |
| $x_{15}$: | Total fertility rate (per woman) |
| $x_{16}$: | Foreign direct investment (PPP USD) |
| $x_{17}$: | Gross fixed capital formation (PPP USD) |
| $x_{18}$: | Trade (PPP USD) |
| For the measurement part | |
| $w_1$: | Control of Corruption |
| $w_2$: | Rule of Law |
| $w_3$: | Population with access to essential drugs (%) |
| $w_4$: | Population using adequate sanitation facilities (%) |
| $w_5$: | Public expenditure on education (% of GDP) |

## Table 1. Results of the Measurement Model

| Dep. var. / Expl. var. | $y_1$ | $y_2$ | $y_3$ | $y_5$ | $y_6$ | $y_7$ | $y_8$ |
|---|---|---|---|---|---|---|---|
| $y_1^*$ | — | — | — | — | 1 (0) | 0.708 (0.056) | — |
| $y_2^*$ | — | — | — | 1 (0) | — | — | −3.865 (0.343) |
| $y_3^*$ | 1 (0) | 0.662 (0.035) | 0.395 (0.019) | — | — | — | — |
| $w_3$ | — | — | — | 0.042 (0.029) | — | — | −0.103 (0.092) |
| $w_5$ | — | — | — | — | 1.719 (0.823) | 1.584 (0.834) | — |
| $R^2$ | 0.921 | 0.880 | 0.951 | 0.834 | 0.868 | 0.796 | 0.969 |

Figures inside parentheses are standard deviations.

**Table 2. Results of the Structural Equation Model**

| Dependent variables / Explanatory variables | $y_1^*$ | $y_2^*$ | $y_3^*$ |
|---|---|---|---|
| $y_1^*$ | — | — | 0.011 (0.004) |
| $y_2^*$ | 1.374 (0.269) | — | — |
| $y_3^*$ | — | 0.284 (0.308) | — |
| $w_1$ | — | — | 0.614 (0.180) |
| $w_4$ | — | 0.065 (0.023) | — |
| $x_7$ | −0.030 (0.005) | — | — |
| $x_{11}$ | −64.293 (30.547) | — | — |
| $x_{12}$ | — | 0.001 (0.006) | — |
| $x_{13}$ | — | — | 0.077 (0.005) |
| $x_{14}$ | 0.584 (0.588) | — | — |
| $x_{15}$ | — | −4.003 (0.481) | — |
| $R^2$ | 0.821 | 0.798 | 0.892 |

Figures inside parentheses are standard deviations.

## Table 3. Explanations of abbreviations used in rank tables

| | |
|---|---|
| hdi: | human development index |
| $\hat{H}$: | our aggregate index based on estimated factor scores |
| GDP: | Gross Domestic Product per capita |
| $y^*1$: | 'knowledge' or 'education' dimension |
| $y^*2$: | 'health' dimension |
| $y^*3$: | 'political freedom' dimension |
| rhdi: | rank according to HDI |
| rhhat: | rank according to $\hat{H}$ |
| rgdpn: | rank according to (normalised) GDP |
| $ry^*n$: | rank according to $y_n^*$ for n=1,2,3 |

## Table 4. Rank Correlations

| rhhat,rhdi | rhav,rhdi | rhdi, ry*1 | rhdi, ry*2 | rhdi, ry*3 | rhdi,rgdpn | rhhat,rgdpn |
|---|---|---|---|---|---|---|
| 0.861 | 0.915 | 0.917 | 0.916 | 0.528 | 0.756 | 0.756 |
| hdi,hhat | hdi,hav | hdi,y*1 | hdi,y*2 | hdi,y*3 | hdi,gdpn | hhat,gdpn |
| 0.851 | 0.914 | 0.948 | 0.94 | 0.426 | 0.891 | 0.796 |

## Table 5. Country Rankings

| COUNTRY | rankhdi | rankhhat | rankgdpn | ranky*1 | ranky*2 | ranky*3 |
|---|---|---|---|---|---|---|
| Argentina | 1 | 2 | 2 | 2 | 3 | 6 |
| Hungary | 2 | 6 | 1 | 5 | 2 | 12 |
| Slovakia | 3 | 5 | 3 | 1 | 4 | 2 |
| Chile | 4 | 3 | 31 | 9 | 5 | 31 |
| Uruguay | 5 | 4 | 6 | 3 | 6 | 5 |
| Costa Rica | 6 | 12 | 12 | 4 | 9 | 3 |
| Mexico | 7 | 8 | 4 | 10 | 1 | 29 |
| Panama | 8 | 9 | 5 | 16 | 12 | 19 |
| Bulgaria | 9 | 19 | 7 | 18 | 21 | 8 |
| Romania | 10 | 1 | 15 | 17 | 19 | 4 |
| Colombia | 11 | 10 | 11 | 23 | 8 | 9 |
| Mauritius | 12 | 23 | 23 | 8 | 10 | 35 |
| Venezuela | 13 | 16 | 9 | 11 | 23 | 10 |
| Thailand | 14 | 29 | 20 | 21 | 13 | 23 |
| Brazil | 15 | 14 | 14 | 6 | 22 | 1 |
| Philippines | 16 | 15 | 30 | 15 | 7 | 16 |
| Kazakhstan | 17 | 31 | 8 | 7 | 11 | 51 |
| Peru | 18 | 7 | 17 | 29 | 14 | 14 |
| Jamaica | 19 | 21 | 26 | 25 | 15 | 15 |
| Turkey | 20 | 13 | 13 | 14 | 16 | 54 |
| Sri Lanka | 21 | 37 | 10 | 34 | 32 | 47 |
| Paraguay | 22 | 22 | 22 | 22 | 33 | 37 |
| Dominican Rep. | 23 | 35 | 18 | 31 | 25 | 7 |
| Uzbekistan | 24 | 11 | 29 | 13 | 37 | 41 |
| China | 25 | 18 | 38 | 28 | 27 | 50 |
| Iran | 26 | 27 | 27 | 33 | 36 | 38 |
| Jordan | 27 | 34 | 16 | 20 | 34 | 13 |
| Kyrgyzstan | 28 | 38 | 25 | 24 | 18 | 18 |
| Guyana | 29 | 20 | 19 | 26 | 26 | 27 |
| Algeria | 30 | 41 | 36 | 36 | 20 | 46 |

## Table 5. Country Rankings: contd.

| COUNTRY | rankhdi | rankhhat | rankgdpn | ranky*1 | ranky*2 | ranky*3 |
|---|---|---|---|---|---|---|
| South Africa | 31 | 39 | 39 | 35 | 39 | 21 |
| Syria | 32 | 36 | 21 | 12 | 24 | 22 |
| Vietnam | 33 | 28 | 32 | 19 | 29 | 34 |
| Indonesia | 34 | 26 | 34 | 37 | 28 | 11 |
| Bolivia | 35 | 47 | 37 | 27 | 38 | 39 |
| Egypt | 36 | 17 | 28 | 32 | 30 | 49 |
| Honduras | 37 | 46 | 40 | 30 | 17 | 20 |
| Guatemala | 38 | 25 | 24 | 40 | 46 | 28 |
| Morocco | 39 | 51 | 35 | 38 | 41 | 42 |
| Zimbabwe | 40 | 33 | 41 | 42 | 31 | 48 |
| Ghana | 41 | 32 | 33 | 43 | 35 | 36 |
| Cambodia | 42 | 30 | 53 | 49 | 43 | 45 |
| Kenya | 43 | 24 | 52 | 41 | 40 | 43 |
| Pakistan | 44 | 50 | 44 | 47 | 45 | 17 |
| Togo | 45 | 43 | 48 | 39 | 50 | 26 |
| Bangladesh | 46 | 40 | 42 | 45 | 44 | 55 |
| Madagascar | 47 | 42 | 45 | 44 | 42 | 40 |
| Mauritania | 48 | 45 | 46 | 51 | 47 | 52 |
| Zambia | 49 | 49 | 50 | 50 | 53 | 53 |
| Senegal | 50 | 54 | 55 | 46 | 51 | 30 |
| Benin | 51 | 44 | 43 | 53 | 49 | 44 |
| Guinea | 52 | 53 | 51 | 55 | 52 | 25 |
| Gambia | 53 | 48 | 47 | 52 | 55 | 24 |
| Mali | 54 | 55 | 54 | 48 | 48 | 32 |
| Chad | 55 | 52 | 49 | 54 | 54 | 33 |

## Table 6. Rank Differences

| COUNTRY | rhdi-rhhat | rgdpn-rhhat | ry*1-rhhat | ry*1-rhdi | ry*3-rhhat | ry*3-rhdi |
|---|---|---|---|---|---|---|
| Argentina | -1 | 0 | 0 | 1 | 4 | 5 |
| Hungary | -4 | -5 | -1 | 3 | 6 | 10 |
| Slovakia | -2 | -2 | -4 | -2 | -3 | -1 |
| Chile | 1 | 28 | 6 | 5 | 28 | 27 |
| Uruguay | 1 | 2 | -1 | -2 | 1 | 0 |
| Costa Rica | -6 | 0 | -8 | -2 | -9 | -3 |
| Mexico | -1 | -4 | 2 | 3 | 21 | 22 |
| Panama | -1 | -4 | 7 | 8 | 10 | 11 |
| Bulgaria | -10 | -12 | -1 | 9 | -11 | -1 |
| Romania | 9 | 14 | 16 | 7 | 3 | -6 |
| Colombia | 1 | 1 | 13 | 12 | -1 | -2 |
| Mauritius | -11 | 0 | -15 | -4 | 12 | 23 |
| Venezuela | -3 | -7 | -5 | -2 | -6 | -3 |
| Thailand | -15 | -9 | -8 | 7 | -6 | 9 |
| Brazil | 1 | 0 | -8 | -9 | -13 | -14 |
| Philippines | 1 | 15 | 0 | -1 | 1 | 0 |
| Kazakhstan | -14 | -23 | -24 | -10 | 20 | 34 |
| Peru | 11 | 10 | 22 | 11 | 7 | -4 |
| Jamaica | -2 | 5 | 4 | 6 | -6 | -4 |
| Turkey | 7 | 0 | 1 | -6 | 41 | 34 |
| Sri Lanka | -16 | -27 | -3 | 13 | 10 | 26 |
| Paraguay | 0 | 0 | 0 | 0 | 15 | 15 |
| Domin. Rep. | -12 | -17 | -4 | 8 | -28 | -16 |
| Uzbekistan | 13 | 18 | 2 | -11 | 30 | 17 |
| China | 7 | 20 | 10 | 3 | 32 | 25 |
| Iran | -1 | 0 | 6 | 7 | 11 | 12 |
| Jordan | -7 | -18 | -14 | -7 | -21 | -14 |
| Kyrgyzstan | -10 | -13 | -14 | -4 | -20 | -10 |
| Guyana | 9 | -1 | 6 | -3 | 7 | -2 |
| Algeria | -11 | -5 | -5 | 6 | 5 | 16 |

## Table 6. Rank Differences: contd.

| COUNTRY | rhdi-rhhat | rgdpn-rhhat | ry*1-rhhat | ry*1-rhdi | ry*3-rhhat | ry*3-rhdi |
|---|---|---|---|---|---|---|
| South Africa | -8 | 0 | -4 | 4 | -18 | -10 |
| Syria | -4 | -15 | -24 | -20 | -14 | -10 |
| Vietnam | 5 | 4 | -9 | -14 | 6 | 1 |
| Indonesia | 8 | 8 | 11 | 3 | -15 | -23 |
| Bolivia | -12 | -10 | -20 | -8 | -8 | 4 |
| Egypt | 19 | 11 | 15 | -4 | 32 | 13 |
| Honduras | -9 | -6 | -16 | -7 | -26 | -17 |
| Guatemala | 13 | -1 | 15 | 2 | 3 | -10 |
| Morocco | -12 | -16 | -13 | -1 | -9 | 3 |
| Zimbabwe | 7 | 8 | 9 | 2 | 15 | 8 |
| Ghana | 9 | 1 | 11 | 2 | 4 | -5 |
| Cambodia | 12 | 23 | 19 | 7 | 15 | 3 |
| Kenya | 19 | 28 | 17 | -2 | 19 | 0 |
| Pakistan | -6 | -6 | -3 | 3 | -33 | -27 |
| Togo | 2 | 5 | -4 | -6 | -17 | -19 |
| Bangladesh | 6 | 2 | 5 | -1 | 15 | 9 |
| Madagascar | 5 | 3 | 2 | -3 | -2 | -7 |
| Mauritania | 3 | 1 | 6 | 3 | 7 | 4 |
| Zambia | 0 | 1 | 1 | 1 | 4 | 4 |
| Senegal | -4 | 1 | -8 | -4 | -24 | -20 |
| Benin | 7 | -1 | 9 | 2 | 0 | -7 |
| Guinea | -1 | -2 | 2 | 3 | -28 | -27 |
| Gambia | 5 | -1 | 4 | -1 | -24 | -29 |
| Mali | -1 | -1 | -7 | -6 | -23 | -22 |
| Chad | 3 | -3 | 2 | -1 | -19 | -22 |

# Conclusions

## On the theoretical side

We have presented a framework that

distinguishes and provides a link between *latent* capabilities and *observed* functionings,

offers an *explanation* of the level of capabilities both in terms of endogenous and exogenous factors,

allows for the possibility to have different functionings for the same level of capabilities due to the presence of exogenous variables,

and can *predict* capability levels.

## On the empirical side

What are the lessons to be learnt from our model results and rank comparisons?

Include as many dimensions as possible while computing any measure of overall development or welfare because each new component does have a significant contribution

A better social and political environment not only helps the 'realisation' of capabilities but also augments the level of capabilities themselves

Thus the State has a role to play and a positive one in terms of better social infrastructure and better governance

To lead the way to a 'virtuous' development cycle