Models of Unobserved Latent Variables
Advanced Econometrics (Subject 310)

Dr. M. Weeks
Faculty of Economics & Politics, University of Cambridge

January 2000

Contents

1. Introduction

Within the social sciences and in particular fields such as economics, education and psychology, many of the phenomenon that we wish to explain are either measured with error or not directly measurable. Other examples include intelligence, education, management expertise and institutional change. Perhaps the most prevalent example in economics is based upon Friedman's (1957) permanent income hypothesis where there is no directly observable measure of permanent income[1]. In practice, the existence of true measurements for theoretical counterparts is the exception to the rule. In such cases a number of alternate approaches have been advocated. In time series analysis Harvey (1981) has developed structural time series models to handle unobserved components such as the business cycle, seasonality and trend. More recently economists have analysed the extent to which monetary policy is endogenously determined. However, given that there is no single measurable quantity which represents *monetary policy*, studies such as Avery (1979) have used a model specification where it assumed that monetary policy is represented by a single *latent* variable, and that policy is manifest in the behaviour of a set of indicators. In Lahiri (1976) the author examines the impact of the Fisher effect of inflationary expectations on the nominal interest rate, where the unobserved price expectations variable is modelled using a structural equations approach. Of relevance to this research Kaufmann, Kraay, and Zoido-Lobaton (1999) adopt a similar approach in examining cross-country variation in three broad areas of governance: probity, bureaucratic quality and the rule of law.

The modelling strategy we will discuss has been extensively used in psychometrics and more recently in econometrics. It is founded upon the specification of a system of equations which specify the relationship between a set of unobservable latent variables, $\mathbf{y}^*$, a set of observable endogenous indicators $\mathbf{y}$, and a set of observable exogenous variables $\mathbf{x}$. This approach builds upon the early work of Joreskog and Goldberger (1975) and Zellner (1970), and has been formalised in the LISREL[2] model of a set of linear structural equations.

2. Regression with an Unobserved Exogenous Variable

Consider the following linear regression model

[1] See Crockett (1960)

[2] LISREL is an acronym for *linear structural relationships.*

$$y = \delta z^* + \varepsilon, \tag{2.1}$$

where y is an observed scalar, z^* is an unobserved regressor and ε is a stochastic error term. z^* is in standardised units (mean zero, variance 1).

The Observed Data

The analyst will only observe the variance and mean of y, where the variance of y is given by $\delta^2 + \sigma_\varepsilon^2$. Obviously δ is not identified. To circumvent this problem we need to introduce additional information. We do this by postulating the existence of a *measurement* equation for z^* which we write as

$$z = z^* + u \tag{2.2}$$
$$z^* = \boldsymbol{\beta}'\mathbf{x}, \tag{2.3}$$

where $\mathbf{x} = (x_1, x_2, ..., x_s)'$ is a $s \times 1$ vector of *causes* and $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_s)'$ is a $s \times 1$ vector of unknown parameters. We assume that $E[\varepsilon u] = 0$.

Equations (2.1), (2.2) and (2.3) constitute the *structural* model.

Example 2.1. *Friedmans (1957) Permanent Income Model*

Think of (2.1) as the consumption function, (2.2) as measured income and (2.3) as the multiple causes of permanent income.

The Reduced Form

If we combine (2.1), (2.2) and (2.3) we may write the reduced form as

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{z} \end{bmatrix} = \begin{bmatrix} \mathbf{x} & 0 \\ 0 & \mathbf{x} \end{bmatrix} \begin{bmatrix} \boldsymbol{\pi} \\ \boldsymbol{\beta} \end{bmatrix} + \begin{bmatrix} \varepsilon \\ u \end{bmatrix} \tag{2.4}$$

where $\boldsymbol{\pi} = \boldsymbol{\delta\beta}$ and Σ_y , the covariance matrix of ε, u, is diagonal.

We note that an alternate approach is based upon the method of instrumental variables. For example, in the context of the above problem we might have available a variable q which is correlated with z^* and y but not with u. It is also worth emphasising that if the model contains a number of fully observed exogenous variables and one explanatory variable which is measured with error, then the following trade-off exists. One strategy would be to simply drop the unobservable regressor and incurr the normal omitted variable bias. Alternately, we could include a proxy, or indicator variable, However, since by definition this variable will contain measurement error, parameter estimates will also be inconsistent. Studies by McCallum (1972) and Wickens (1972) have demonstrated that based upon the criteria of asymptotic bias, the use of a proxy variable is preferred.

2.0.1. Estimation: ML or Two-Step Approach

A Simle Two-step Method

Note that the measurement equation $z = \boldsymbol{\beta}'\mathbf{x} + u$ satisfies all the requirements of OLS. Thus $\boldsymbol{\beta}$ and the asymptotic covariance matrix for $\boldsymbol{\beta}$ can estimated. For the structural equation

$$y = \delta\left(\boldsymbol{\beta}'\mathbf{x}\right) + \varepsilon, \tag{2.5}$$

we may substitute $\hat{z}^* = \hat{\boldsymbol{\beta}}'\mathbf{x}$ and obtain an estimate of δ. Pagan (1984) shows that this procedure is consistent, as asymptotically efficient as full MLE and much easier to implement. Note that we may think of $\hat{z}^*$ as an instrumental variable for the unobserved z^*.

3. Mutiple Indicators of an Unobserved Latent Variable

In section 2 we saw that for y and z^* scalar quantities we cannot estimate the parameters without the use of additional information. Above this information took the form of a measurement equation for the unobserved exogenous variable z^*, which is based upon a model of *multiple causes* of z^*.

Below we consider how information in the form of *multiple indicators* of z^* can also assist in identifying model parameters.

Let $\mathbf{y} = (y_1, y_2, ..., y_m)'$ denote a $m \times 1$ vector of observed indicators

$\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_m)'$ is a $m \times 1$ vector of unknown parameters.

$\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_m)'$ is a $m \times 1$ vector of stochastic terms.

Example 3.1. *m=3*

$$y_j = \beta_j z^* + \varepsilon_j, \; j = 1, 2, 3. \tag{3.1}$$

We let σ_j^2 denote the variance of ε_j and set $cov(\varepsilon_j, \varepsilon_i) = 0 \; \forall i \neq j$. Based upon (3.1) there are six unknown parameters: $\boldsymbol{\theta} = (\beta_1, \beta_2, \beta_3, \sigma_1^2, \sigma_2^2, \sigma_3^2)'$.

What do we observe?

The analyst observes $\Sigma_{y,}$ the 3×3 covariance matrix of $\mathbf{y}$, where the diagonal elements, s_{jj}, represent the variance of y_j, written as

$$s_{jj} = \beta_j^2 + \sigma_j^2$$

The off-diagonal elements, $s_{12,}s_{13,}$ and s_{23} are, respectively, given by $\beta_1\beta_2$, $\beta_1\beta_3$, and $\beta_2\beta_3$.

3.1. Solution

An application of the Method-of-Moments (MOM) allows us to express the unknown population moments, $\boldsymbol{\theta}$, in terms of the six pieces of observed (sample) moments. This sample information is contained in the unique[3] elements of Σ_y.

$$\begin{aligned}
\beta_1 &= \frac{s_{12}s_{13}}{s_{23}}^{0.5} \\
\beta_2 &= \frac{s_{12}s_{23}}{s_{13}}^{0.5} \\
\beta_3 &= \frac{s_{13}s_{23}}{s_{12}}^{0.5} \\
\sigma_j^2 &= s_{jj} - \beta_j^2
\end{aligned}$$

Based upon the moments of the sampled data we have enough information to exactly identify the structural parameters in $\boldsymbol{\theta}$. Note also that if $\boldsymbol{\varepsilon} \sim mvn(\mathbf{0}, \Sigma_y)$ then these estimates are ML with the attendant properties.

Example 3.2. *m=M*

$$\mathbf{y} = \boldsymbol{\beta} y^* + \boldsymbol{\varepsilon} \tag{3.2}$$

$$Var(\mathbf{y}) = \Sigma_y = \boldsymbol{\beta}\boldsymbol{\beta}' + \Theta^2 \tag{3.3}$$

where Θ^2 is a diagonal matrix with elements σ_j^2.

If $M > 3$ there exists more unique sample variance and covariance terms in Σ_y than the $2M$ unknown parameters. For $M < 3$ the model is underidentified.

Multiple indicators provide one method for identifying population parameters and is the same route that is used in factor analysis.

[3]Note: for a $m \times m$ covariance matrix there are $m(m+1)/2$ unique elements.

4. Multiple Indicators and Multiple Causes

The canonical form of a structural equations model in which we combine the models of the last two sections is given by the following system of linear structural relations

$$\mathbf{y}^* = \boldsymbol{\beta}\mathbf{y}^* + \boldsymbol{\gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}, \tag{4.1}$$

where $\mathbf{y}^* = (y_1^*, y_2^*, ..., y_l^*)'$ and $\boldsymbol{\xi} = (\xi_1, \xi_2, ..., \xi_n)'$ are, respectively, $l \times 1$ and $n \times 1$ unobservable random vectors of latent dependent and independent variables. $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are, respectively, $l \times l$ and $l \times n$ matrices of unknown parameters, $\boldsymbol{\zeta} = (\zeta_1, \zeta_2, ..., \zeta_l)'$ is a vector of structural disturbance terms. The elements of $\boldsymbol{\beta}$ represent the direct influence of latent variables on other latent variables in the system. Elements of $\boldsymbol{\gamma}$ represent the influence of latent explanatory variables on $\mathbf{y}^*$. We assume that $\boldsymbol{\zeta}$ is uncorrelated with $\boldsymbol{\xi}$. Vectors of indicators for $\mathbf{y}^*$ and $\boldsymbol{\xi}$ are given, respectively, by $\mathbf{y} = (y_1, y_2, ..., y_m)'$ and $\mathbf{x} = (x_1, x_2, ..., x_s)'$, with the two measurement equations given by

$$\mathbf{y} = \Lambda^y \mathbf{y}^* + \boldsymbol{\varepsilon} \tag{4.2}$$

and

$$\mathbf{x} = \Lambda^x \boldsymbol{\xi} + \boldsymbol{\delta}. \tag{4.3}$$

Λ^y and Λ^x denote $m \times l$ and $s \times n$ parameter matrices. $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$ represent errors of measurement which may exhibit intra-set correlation. We assume that $cov(\varepsilon_j, \varepsilon_i) = 0 \ \forall i \neq j$ such that any correlation across the indicators is driven by the common factor y^*. We let Θ_ε and Θ_δ represent the covariance matrices of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$. (4.1) represents the most general form of the LISREL formulation. Note that we may think of (4.1) and (4.2) as factor analysis models for the observable variables $\mathbf{y}$ and $\mathbf{x}$, where Λ^y and Λ^x are the factor loadings.To specify a model in which $\boldsymbol{\xi}$ is measured without error one simply sets $\Lambda^x = \mathbf{I}_x$, where $\mathbf{I}_x$ is a $s \times s$ identity matrix.

We might also think of elements of $\mathbf{y}$ as possible instruments for y^*. For example, if only a single instrument is available an IV approach could be utilised. Faced with multiple instruments that are correlated, then one or more elements of $\mathbf{y}$ could be selected using prior information. Alternately principle component analysis (PCA) could be used to construct a synthetic variable. Following this (4.1) could be estimated using either an instrument or a synthetic variable in place of the unobserved y^*. The major differences between this approach and the MIMIC are two fold. First, the MIMIC model utilises the information contained in *all* the multiple indicators in $\mathbf{y}$, and therefore will represent higher efficiency relative to either IV or PCA. Second, the combination of both a structural equation and a measurement equation[4].

5. A Structural Equations Approach to Modelling Institutional Change

An innovative application of this modelling strategy is to determine empirically the factors behind the demand for institutional change. This approach is founded upon the belief that no single observed variable can adequately measure institutional change and therefore a latent variable representation is utilised which explicitly accounts for measurement error. Indicators of institutional change are corporate governance, competition policy, financial regulation and EBRD legal transition indicators.[5] The causes include budget constraints based upon explicit subsidies and tax data, external factors influencing institutional reform and an index of liberalisation.[6] These variables are listed in Table 1.1. The principal advantage of this approach is that it does not rely on exact measurement of institutions. In addition, using an estimate of the variance of the stochastic term for each indicator, estimates of how informative each indicator is with regards to institutional change can be generated. A welcome side-product of our approach is that it is possible to create an index of institutional change based on the observable variation in both EBRD's transition indicators and the exogenous variables, which may be represented as a weighted average of the former, with endogenously generated weights.

The structure of the model is as follows. Each y_i $(i = 1, ..., m)$ represents an independent indicator of institutional change, denoted y^*, such that we may write

$$y_j = \Lambda_j^y y^* + \varepsilon_j,\ j = 1, \ldots, m. \tag{5.1}$$

[4]Seminal studies using the MIMIC model include Zellner (1970),Hauser and Goldberger (1971), and Goldberger (1972).

[5]For a complete discussion of the legal transition indicators see Annex 2.2 of the 1999 EBRD Transition Report.

[6]See Section * for a full explanation of the indicators and causes. (Maria: *we need to summarise Annex 2.2 here for EBRD legal transition indicators*).

We ley $\boldsymbol{\tau}$ denote the $m \times 1$ vector of diagonal elements of Θ_ε. We also posit that the institutional change is linearly determined by a set of observable exogenous variables $\mathbf{x}$, subject to an error ζ, giving

$$y^* = \alpha + \gamma_1 x_1 + \gamma_2 x_2 + ... + \gamma_s x_s + \zeta. \tag{5.2}$$

Examining (5.1) and (5.2) we may think of the model as comprising two parts: (5.2) is the structural (or state) equation and (5.1) is the measurement equation reflecting that the observed measurements are imperfect indicators. The measurement equation specifies how the unobservable construct institutional change is determined by the observed endogenous variables and the structural equation specifies the causal relationship between the observed exogenous causes and institutional change[7,8]. In this instance (5.2) is a special case of a factor analysis model with a set of l observable indicators determined linearly by a single unobserved (common) factor - institutional change. Figure 1 adapted from Chen (1981) graphically illustrates the relationship between the indicators and causes of institutional change using a path diagram..

The two key assumptions which underly our modelling strategy are that: i) measurement errors in the individual indicators of institutional change are uncorrelated across indicators namely $cov(\varepsilon_i, \varepsilon_j) \neq 0 \ \forall_i = j$; and 2) the relationship between unobserved institutional change and the observed indicators is linear. One way of circumventing the second assumption has been proposed by Kaufmann, Kraay, and Zoido-Lobaton (1999) and simply forms a composite indicator by aggregating over the observed ordinal indicators. Why does this not require linearity assumption?

Since y^* is unobserved it is not possible to recover direct estimates of $\boldsymbol{\gamma}$. However if we combine these two equations and solve for the reduced form representation, then based upon a sample of T observations we may write

$$\mathbf{y} = \mathbf{x}\boldsymbol{\pi} + \mathbf{v}, \tag{5.3}$$

where $\boldsymbol{\pi} = \Lambda^y \boldsymbol{\gamma}'$ is the $m \times s$ reduced form coefficient matrix and $\mathbf{v} = \Lambda^y \zeta + \boldsymbol{\varepsilon}$ is the reduced form disturbance with covariance matrix

$$\Theta_v = E[(\Lambda^y \zeta + \boldsymbol{\varepsilon})(\Lambda^y \zeta + \boldsymbol{\varepsilon})' = \sigma_\zeta^2 \Lambda^y \Lambda^{y\prime} + \Theta_\varepsilon. \tag{5.4}$$

[7] More general forms of (**??**) and (5.2) are possible including models which specify a measurement equation for $\mathbf{x}$ and models which allow elements of $\mathbf{y}^*$ to appear on the RHS of (5.2).

[8] See Table 1.2 for dimension of parameter matrices.

Note that the structure of the reduced form covaraince matrix Θ_v is characteristic of factor analysis models where the correlations between the observed variables (here indicators) are accounted for by the unobserved (common) latent variable. In this instance the common factor is ζ, $\boldsymbol{\varepsilon}$ denotest the unique factor, and Λ^y the vector of factor loadings.

Based upon equations (5.3) and (**??**) there are two sets of restrictions on the reduced form[9]. First, the $m \times s$ coefficient matrix $\boldsymbol{\pi}$ has rank 1, since the ms elements of $\boldsymbol{\pi}$ are expressed in terms of the $m + s$ elements of Λ^y and $\boldsymbol{\gamma}$. Second, the $m \times m$ covariance matrix Θ_v represents the sum of a rank one matrix and a diagonal matrix, $\Theta_{\boldsymbol{\varepsilon}}$. The $m(m+1)/2$ unique elements of Θ_v are expressed in terms of the $1 + 2m$ elements of Λ^y, σ^2_ζ, and $\boldsymbol{\tau}$.

The question of identification can be addressed by examining equations (5.2) and (5.4). Here we see that the relationship between observable moments and structural parameters may be written as

$$E[\mathbf{yy}'] = \sigma^2_\zeta \Lambda^y \Lambda^{y\prime} + \Theta_\varepsilon \tag{5.5}$$
$$E[\mathbf{xy}'] = E[\mathbf{xx}'] \Lambda^y \boldsymbol{\gamma}' \tag{5.6}$$

which using may be trivially rewritten in terms of the reduced form parameters, namely,

$$\boldsymbol{\pi} = \Lambda^y \boldsymbol{\gamma}' = E[\mathbf{xx}']^{-1} E[\mathbf{xy}'].$$

Since equation (5.6) expresses the $q = ms$ observable moments in terms of the $p = m + s$ structural paramters, then if $q - p \geq 0$ the set of mean paramters will be identified. If this condition holds, then the remaining paramters in equation (5.5) will be identified.

The existence of an unobserved latent variable y^* provides for a complex reduced form where the vector of parameters Λ^y appears in both the reduced form coefficient matrix and the covariance matrix. As noted by Chen (1981) this makes the method of ML intractable. If y^* were observed then each equation could be estimated separately using OLS. Exploiting this fact the parameters of (5.1) and (5.2) may be estimated using the EM algorithm which circumvents this problem

[9] As presented (5.3) and (**??**) are indeterminate since the reduced form parameters are invariant to a transformation given by $\Lambda^y \varpi$, $\boldsymbol{\gamma}/\boldsymbol{\varpi}$ and σ^2_ζ/ϖ, where ϖ is a scalar. The normalisation $\sigma^2_\zeta = 1$ or setting one element of $\Lambda^y = 1$ circumvents this problem.

by working with a set of initial estimates of $\boldsymbol{\gamma}$ and Λ^y and an estimate of the conditional moment $E[y^*|\mathbf{X}, \mathbf{y}]$ which is given by

$$\hat{y}_c^* = \hat{E}[y^*|\mathbf{X}, \mathbf{y}] = (1 + \hat{\Lambda}\hat{\Theta}_v\hat{\Lambda}^y)^{-1}\mathbf{X}\hat{\boldsymbol{\gamma}} + \hat{\boldsymbol{\gamma}}\hat{\Theta}_v^{-1}\hat{\Lambda}^y.$$

For example, with $\hat{y}_c^*$ denoting an estimate of the conditional expectation of y^*then updated estimates of $\boldsymbol{\gamma}$ together with variances σ_ζ^2 and $\sigma_{\varepsilon_j}^2$, $j = 1, \ldots, m$ can be obtained from the standard least squares formulae

$$\begin{aligned} \boldsymbol{\gamma} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}^* && (5.7) \\ \Lambda^y &= (\mathbf{y}^{*\prime}\mathbf{y}^*)^{-1}\mathbf{y}^{*\prime}\mathbf{y}, && (5.8) \end{aligned}$$

where $\mathbf{X}$, $\mathbf{y}^*$ and $\mathbf{y}$ are data matrices.

6. A MIMIC Model of Social Status and Participation

Hodge and treiman (1968) examined the relationship between social status and particiaption in a sample of 530 women. The model may be written as

$$y^* = \gamma_1 x_1 + \gamma_2 x_2 + \gamma_3 x_3 + \zeta \quad (6.1)$$

$$y_1 = \lambda_1 y^* + \varepsilon_1,\ y_2 = \lambda_2 y^* + \varepsilon_2,\ y_3 = \lambda_3 y^* + \varepsilon_3 \quad (6.2)$$

where 6.1 denotes the structural equation for the unobserved construct social participation (y^*) with x_1, x_2 and x_3, representing three causes of participation, respectively, income, occupation and education. 6.2 denotes the set of indicator equations for y^*, with y_1 (church attendance), y_2 (memberships) and y_3 (friends see), denoting indicators of participation. A path diagram which summarises this model is given in figure 2.

References

Avery, R. B. (1979): "Modeling Monetary Policy as an Unobserved Variable," *Journal of Econometrics*, 10, 291–311.

Chen, C.-F. (1981): "The EM Approach to the Multiple Indicators and Multiple Causes Model Via the Estimation of the Latent Variable," *Journal of the American Statistical Association*, 76(375).

Goldberger, A. S. (1972): "Maximum-Likelihood Estimation of Regressions Containing Unobservable Independent Variables," *International Economic Review*, 13, 1–15.

Harvey, A. (1981): *Time Series Models.* (Halsted Press, New York).

Hauser, R. M., and A. S. Goldberger (1971): "The Treatment of Unobservable Variable in Path Analysis," in *Sociological Methodology*, ed. by H. L. Costner, pp. 81–177. Jossey-Bass, San Francisco.

Joreskog, K., and A. Goldberger (1975): "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable," *Journal of the American Statistical Association*, 70, 631–639.

Kaufmann, D., A. Kraay, and P. Zoido-Lobaton (1999): "Aggregating Governance Indicators," World Bank, mimeo.

Lahiri, K. (1976): "Inflationary Expectations: Their Formation and Interest Rate Effects," *American Economic Review*, 66, 124–131.

Zellner, A. (1970): "Estimation of Regression Relationships Containing Unobservable Variables," *International Economic Review*, 11, 441–454.

Schematic Representation of MIMIC Model of Social Participation

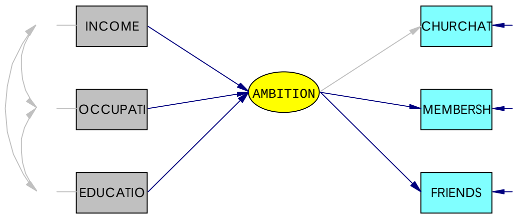

Figure 6.1: Schematic Representation of Mimic Model

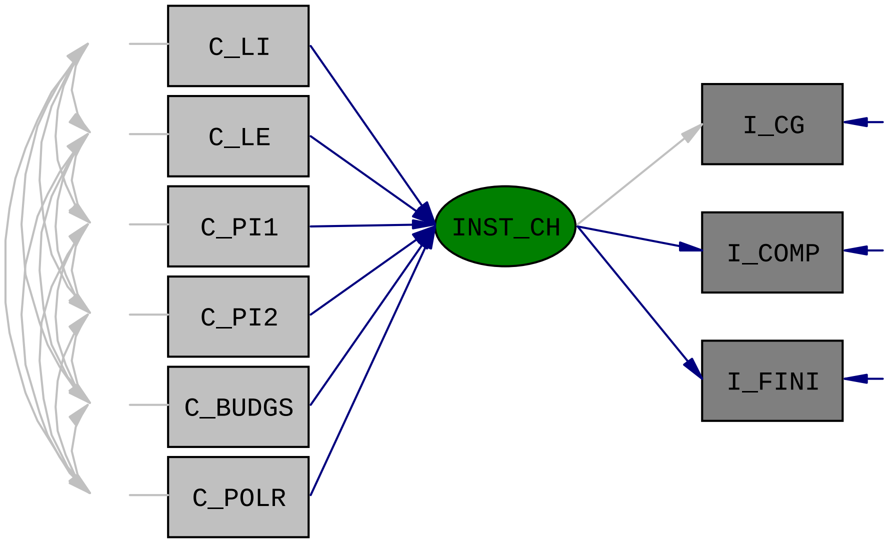

Table 1.1: Multiple Indicators and Multiple Causes of Institutional Change

Causes	Variable Name
Reform Measures	
WB measures	
Liberalisation	
WB Internal Lib. Index	WB_IL
WB External Lib. Index	WB_EL
Privatisation	
WB Privatisation Index	WB_P
EBRD Measures	
Small Scale Privatisation	SCP
Price Liberalisation	PL
Trade	
Share of Trade with EU	EX_shE
Share of Trade with ROW	EX_row
Political Factors	
Civil Liberties	civilib
Political Reform	polright
Freedom Rating	freedom
Macro Causes	
No. years inflation <30%; budget deficit <5%	yearsIBD
years since macroeconomic stabilisation	yearsMS
Fiscal	
Budgetary Subsidies (% GDP)	BUDG_S (242)
Indicators:	
*EBRD Transition Indicators**	
Enterprise Restructuring	ER_ST (267)
Competition Policy	COM_POL (272)
Banking Reform and Interest Rate	BR_I (275)
Overall Legal Extensiveness...	O_LE
Securities and Non-Bank	S_NB
Alternative Measures	
Corporate Governance	CORPG (133)
Competition Policy Index	COMP_I (137)
Rule of Law	ROL

* EBRD Transition Indicators are all measured on a common scale