Applying General to Specific Methodology for Panel Studies of Sources of Economic Growth in Transition Economies

by

Andrii Oliinyk

A thesis submitted in partial fulfillment of
the requirements for the degree of

Master of Arts in Economics

National University of "Kyiv-Mohyla Academy"
Economics Education and Research Consortium
Master's Program in Economics

2003

Approved by ____
Ms.Svitlana Budagovska (Head of the State Examination Committee)

____

____

____

Program Authorized
to Offer Degree Master's Program in Economics, NaUKMA

Date

National University of "Kyiv-Mohyla Academy"

Abstract

Applying General to Specific Methodology for Panel Studies of Sources of Economic Growth in Transition Economies

by Andrii Oliinyk

Chairperson of the Supervisory Committee: Ms.Svitlana Budagovska,
Economist, World Bank of Ukraine

This thesis is a contribution to economic growth research. Its purpose is to use a mechanized variant of general-to-specific specification searching to the question of what variables to include in the analysis of growth in transition economies. The analysis is conducted on longitudinal data for 27 transition economies for the period of ten years. Out of more than 70 candidate variables are physical and human capital investment; governance quality; international aid flow; population growth; and growth of exports share in GDP clamed to have significantly influenced economic growth in the CEU, Baltic and FSU countries. The research ends up in the fixed effects regression equation, pointing most rewarding reforms direction.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF APPENDIX TABLES

Acknowledgments

I want to express my heartiest gratitude to everybody helping me in writing this thesis. I am obliged to my thesis supervisor professor Peter Kennedy for valuable comments along the way of doing the work, for his ever being exact and prompt. I appreciate the attention and a helpful suggestions dr. Tom Coupe gave me on the part of my thesis work. Many thanks to my friends and family for their moral support and belief in me. Special thanks to my sister, Olena, helping me, besides, with programming.

GLOSSARY

CIS. Commonwealth of Independent States;

CEU. Central and Eastern European (countries);

GSM. General-to-specific methodology: the way of specification searching starting from an extended specification and coming through eliminating insignificant variables and testing to a more parsimonious 'specific' regression.

LSDV. Least Squares Dummy Variable estimator.

GDP. Gross domestic product.

Chapter 1

INTRODUCTION

Economic growth has received considerable attention in theoretical and empirical research literature. Still, there is no commonly agreed-upon model applicable for investigating growth. The most dispute arises upon the question of what explanatory variables should be considered. This is particularly the case in regard to transition economies. The purpose of this thesis is to adjust Hoover and Perez's (2000, 2001) variant of general-to-specific specification searching to the question of what variables to include in the analysis of growth in transition economies.

The issue of economic growth has always been the field of special interest in Economics. And today it remains one of the most speculated and written-upon topics. This bewildering number of attempts in the field resulting in great many various empirical econometric models does not seem to have lead to any universal specification, which would be considered valid beyond any particular model.

Most of the models built so far have been devoted to investigating some specific factor of economic growth by testing its significance in the specification. Starting from Levine and Renelt (1993) a new generation of multi-equation simulation models appeared. This was devoted to testing significance of all possible variables in all possible and impossible combinations. Hoover and Perez's (2000, 2001) general-to-specific methodology is one of the most successful attempts belonging to this generation.

Principal idea behind the GSM consists in moving from the "general" extensive specification with too many explanatory variables towards the "specific" - parsimonious one at each step eliminating insignificant regressors. Hoover and Perez suggest a mechanized version of the methodology. Here out of the total number of more than fifty potential regressors all possible combinations of a far smaller number of variables (seven or fourteen) are chosen one at a time. Then composed of such sets "general" specifications are repeatedly tested and gradually reduced to "terminal" specifications, out of which through a series of encompassing tests the final "specific" regression equation is chosen. Another Hoover and Perez's innovation is splitting the sample into two sub-samples covering the first and the last ninety percent of the sample and requiring the "terminal" specification to have statistically equal coefficients for the both sub-samples.

Hoover and Perez (2000, 2001) applied their method to several data sets. They used samples of more than 100 countries and time-averages for more than 25 years for each country, which constituted their cross-section sets. Countries in transition, certainly, were not included in the samples (the last year in the time span for constructing the averages was 1985). And afterwards this approach has not been applied to investigating growth in transition economies, nor, to my knowledge, to panel data estimation.

In this thesis I stick to the general principles of the Hoover and Perez method. Still, the field of application is rather different: transition economies are supposed to possess specific features of their own (practically showing up in several peculiar explanatory variables), and, besides, panel models are to be estimated differently.

In my model I start from a combination of four variables out of the total number of about seventy. Such combination composes an initial "general" regression (actually – two regressions for the two sub-samples). Then by applying a series of tests it is reduced to the "terminal" specification – specific for this "general" regression. The "terminal" specification is a valid restriction of the "general" regression, can pass all the tests imposed, has all coefficients significant, and has statistically equal coefficients for both sub-samples. Final "specific" regression is found by running a series of encompassing tests on all the "terminal" specifications. The method of estimation chosen is fixed effects for unbalanced sample. Finally, however, to distinguish between the impacts of various dummies, such as initial conditions proxies, the model is reestimated applying the random effects estimator.

Though the method pursued is actually a well-organized variant of data mining, it is generally justified by the fact of uncertainty about specification equation and has a lot of valuable features and implications. It ends in a concrete regression equation, and has, as shown by Hoover and Perez (2000, 2001), very good properties in the balance between committing type 1 and type 2 errors.

Practical implications of the work mainly consist in summarizing the experience of transition economies in overcoming inauspicious heritage and often inadequate initial policies along their economic development paths. Besides, the work brings about the possibility to make inferences on directions of the most effective policies that would spur economic growth in terms of Ukraine or any other transition country.

The rest of the thesis is organized as follows. In Chapter II I present an overview of empirical growth literature. Chapter III summarizes possible general and transition specific factors, ever claimed significantly affecting economic growth, and analyzes data available for their proxies. Chapter IV gives the assumptions and description of the basic model. The results of simulations and further respecifications are presented in chapter V along with conclusions.

Chapter 2

GROWTH STUDIES REVIEW

1. Revising the experience of economic growth modeling.

In the recent decades growth has been the imperishable topic of interest in innumerous economic endeavors. It has eventually become the corner stone of the very science of Economics. A great number of simple articles, theoretical and research papers along with some scholarly books have been devoted to the subject. Resolving the task of explaining economic growth remains a challenge for many prominent economists of today. On the one hand, this interest is justified by those brilliant prospects high growth would offer in rising living standards, economic development, or overcoming poverty. On the other hand, this is the evidence of the absence of the firm theoretical and empirical framework in the field.

All the empirical models of economic growth generally root in the CES production function, and most commonly – in its Cobb-Douglas simplification. Duffy and Papageorgiou (2000) give the reasoning for such simplification. It is the only linearly homogenous production function with a constant elasticity of substitution, where each factor's share of income is constant over time. This is consistent with one of Kaldor's (1961) "stylized facts" of growth. Besides, the specification possesses the properties of linear homogeneity and constant elasticity of substitution.

Solow model, employing Cobb-Douglas production function, may be written or naturally extended to the following possible specifications:

$$\frac{\frac{dy}{dt}}{y} = \alpha + \beta \frac{\frac{dk}{dt}}{k} + \gamma \frac{\frac{dh}{dt}}{h} + \theta Z \,, \tag{1}$$

$$\frac{\frac{dy}{dt}}{y} = \alpha + \beta \frac{y}{k} + \gamma \frac{y}{h} + \delta \frac{\frac{dL}{dt}}{L} + \theta Z \,, \tag{2}$$

$$\frac{\frac{dy}{dt}}{y} = \alpha - \beta y + \gamma h + \gamma \frac{\frac{dh}{dt}}{h} + \theta Z \,,\text{[1]} \tag{3}$$

Here y is the national production per worker (so the regressant is its growth); k – amount of capital per worker; h – human capital per worker. Z is a number of variables to account for the residual growth (not explained by other variables present in the models. θ is the vector of coefficients for Z. Usual modifications are using per capita terms instead of those per worker, and employing further assumption of equal rates of labor force and population growth.

Other specifications may well be possible in the same framework. Those presented here have been the most commonly accepted in empirical studies of economic growth.

Even at first glance the specifications suggest a number of potential problems. The most obvious one is that of missing variables. Missing variables bias is a grave problem for the studies that concentrate on some particular (possible) factor of growth. Z is often the only variable (or a very limited number of variables) of interest in a particular research, aiming at claiming correlation (or the absence of such) between the long-run growth and this variable (or variables).

[1] Sachs and Warner, 1997

This may cover variables of quite different kinds, including those having little or no economic sense, like, for example, regional dummies called to account for missing variables. There are plenty of such studies, where the generally accepted variables of capital accumulation and initial conditions are necessarily included plus one or several variables in addition. The results obtained in such regressions can hardly be regarded valid.

Besides, it is seen that the specification would employ either a measure of initial conditions (as expressed by y or y/k etc.) or a measure of investment (here (dk/dt)/k). Though, many researchers construct their tested specifications accounting for these factors simultaneously (e.g. Levine and Renelt, 1992). Such estimation is bound to generate multicollinearity, and so would end up claiming insignificance of one of the factors.

Further, any of the specifications above would potentially bear the endogeneity problem. Variables in models where structural form regressions are estimated (as opposed to reduced form equations) are bound to be subject to reverse causality (here the example is growth and investment). As to reduced form growth regressions (which are often assumed by researchers) the measures employed there potentially have rather different channels of impact on growth; they are anyway highly correlated and often endogenous (Aron, 1997).

On the whole, empirical growth literature, abounding in studies devoted to investigating some particular policy, environment or other a priori component of economic growth, gives a picture of often contradictory unsystematic research burdened with a whole bunch of serious problems – of data, methodology and identification, which many authors underestimate or choose to ignore (Temple, 1996). This makes it hard to confer any valid inferences about one or another factor. Despite various studies establish statistically significant relationships, they frequently do not test the sensitivity of their results to model specification, outliers, measurement error, endogeneity of regressors and omitted variables bias (Aron, 1997).

There have been several studies suggesting a solution or a partial solution to these problems. One such attempt is Aron (1997), who suggests conducting estimation of the equations as a system. Having his base in North's (1990) institutional framework, he proposes such a system. He distinguishes between the variables, which have direct ways to influence long-run growth, in his study they are economic and institutional variables, and other related variables, like various political measures, which affect growth through the intermediary of institutional and economic variables.

As a way to improve the quality of estimation, panel data models are repeatedly suggested to estimate. The studies conducted in the field so far have most commonly been done on cross-sectional data. The cross-section study, rather than the panel model, is often gravely reasoned by the absence of some time series, like many proxies for human capital (accumulation). But it obviously employs a number of strict assumptions (such as equal cross-country effects from all the variables employed), it leads to less efficient (if consistent) estimates, and is much more vulnerable to omitted variables bias. The potential shortcomings of longitudinal estimation consist in the fact that using the fixed effect estimator in an instrumental variable framework would lead to inconsistent estimates for all but strictly exogenous variables chosen as instruments (Caselli et al, 1996).

One powerful tool to investigate robustness of a factor's influence on growth is applying multistage simulation models, described in much detail in the following section.

2. Looking through the experience of econometric simulations in growth research.

While the most common in the growth research literature are various studies determining relations between some particular factor and growth, there have been several attempts to estimate the validity of these models and create an exhaustible picture of all “true” sources of economic growth. They have been the extreme boundaries sensitivity analysis (EBA, Levine and Renelt, 1992), Bayesian model averaging (BMA, Sala-i-Martin, 1997 a, b), Bayesian average of classical estimates (BASE, Doppelhofer et al., 2000), and general-to-specific methodology (GSM, Hoover and Perez, 2000 and 2001). All the studies investigated well over 50 factor variables that had ever been claimed significant in cross-countries growth regressions.

In 1992 Levine and Renelt founded a new generation of growth models. In their milestone paper "A Sensitivity Analysis of Cross-Country Growth Regressions" they applied Leamer’s (1985) extreme boundaries analysis for cross-section averages or point estimates for 116 countries, 1960 through 1989. The researches tested the robustness of correlation of the per capita growth (the regressant) and any particular regressor, the latter combined with six of all other independent variables (of which proxies for initial conditions and human and physical capital accumulation are fixed and the two others are random). From the possible combinations correlated variables (like different proxies for the same factor) were excluded. If a variable in one of all the regressions was insignificant or changed its sign, it was claimed producing no robust influence upon growth.

No wonder the research finally reached a dismal conclusion and turned a severe criticism of the approach used in cross-country growth regressions, claiming that almost all significant correlations possibly found there do not stand the EBA and the only correlation able to pass its requirements is the one between growth and investment share in GDP. Thus Levine and Renelt claimed that almost everyone indicator considered is fragile: small alterations in the other explanatory variables overturn past results, which implies that there is not a reliable independent statistical relationship between a wide variety of macroeconomic indicators and growth.

The apparent drawback of the investigation is little account for endogeneity. Invariably accounting for investment, the researchers left the other variables the modest channels of influencing growth – through improving efficiency of resource allocation or technological progress. Investment share of GDP was also subjected to the EBA in the paper, and it was found to stand the criteria applied for only its correlation with ratio of international trade to GDP.

The approach was criticized for employing too strict criteria of robustness (Sala-i-Martin (1997 a, b), Doppelhofer et al. (2000)). Things of that kind are likely to happen in a big number of regressions, and a possible way round suggested by Doppelhofer, et al. (2000) is reasonable extreme bounds of Granger and Uhlid (1990), or another one – as in Sala-i-Martin (1997).

Sala-i-Martin (1997 a, b) did his research on a similar data as Levine and Renelt (1992), sticking to their specification, and came to a quite different and much more optimistic conclusion. He constructed weighted averages of all estimates with weights proportional to values of likelihood function (and unweighted as well – to account for possible endogeneity problem). As a measure of significance he built a likelihood-weighted sum of normal cumulative distribution functions. He ultimately assigned the level of significance to each of the variables, which would reflect the probability of the true value of a particular parameter to lie on the one side of zero. Among the 95 percent significant variables were regional variables and absolute latitude (far away from the equator is good for growth); political variables; religious variables; market distortions and market performance; types of investment; primary sector production; openness; type of economic organization.

Doppelhofer, Miller, and Xavier Sala-i-Martin (2000), which study was a kind of extension of the previous work, investigated robustness of correlation for 32 possible explanatory variables for the cross-section of 98 countries, 1960 – 1992. Their method employs diffuse prior with weights applied to different models proportional to values of log-likelihood function for different specifications corrected for degrees of freedom.

Ultimately, the researchers came to the division among the explanatory variables employed distinguishing between the group of four variables having good explanatory power and estimated precisely; a second group, including seven variables, being of a weaker explanatory power, but still having precisely estimated coefficients; and another five variables being marginal: they are claimed to be reasonable regressors if a researcher has a strong prior belief in their relevance. The other sixteen variables have weak explanatory power and are imprecisely estimated.

In the first group of variables there are Sachs and Warner index of openness (this captures tariff- and non-tariff barriers, black-market premium, degree of socialism and monopolization of exports by the government), and, besides, level of income in the initial year 1960 and fraction of GDP in mining (capturing availability of natural resources). The second group comprises life expectancy, the primary school enrolment rate in 1960 (the basic level of investment in human resources); local dummies; the fraction of primary exports in total exports; the real exchange rate distortions. Among variables marginally related to growth there is measure of outward orientation (with negative! partial correlation and average significance of more than 90 percent).

Hoover and Perez (2000 and a bit modified study of 2001) evaluated Levine and Renelt's EBA and Sala-I-Martin's BMA methodologies. Using simulations Hoover and Perez showed that as far as the EBA does perfectly well in avoiding type-1 error, it is apt to commit type-2 error too often. BMA was shown to be too lax: it would almost never reject a true regressor, but can accept a false one.

Instead they suggested using GSM. Starting with including all possible variables into the regression equation, they reduced it to the specific regression. This specific regression was regarded acceptable if it has the properties of being statistically well specified (it stands the tests conducted); being a valid restriction of the general regression, and encompassing every other parsimonious regression that is a valid restriction of the general regression. The researchers used the LSE approach of encompassing regression.

According to the criteria chosen (the size ratio and the power ratio) and in comparison with the EBA and BMA the it was shown that the general to specific "not only usually finds the truth nearly as well as one would if God had whispered the true specification in one's ear, but it also is able to discriminate between true and false variables extremely well".

Hoover and Perez (2000) found for Levine and Renelt's data set: Real per capita GDP in 1960 has a small negative effect on growth; the trade variables appeared to have larger effects. The import share has a large negative effect; the growth of imports has a large positive effect. For Sala-i-Martin's data set the following variables prove important: GDP per capita in 1960; five religious indicators; three financial measures; the standard deviation of inflation (related positively! to growth); investment (only equipment investment); political variables; the rule of law.

The GSM method employed may rise criticism in the part that it can be regarded a sophisticated form of data mining, and besides it produced some awkward result in the part of inflation (the drawback eliminated in the modified study of Hoover and Perez, 2001). However, it produces a valuable result in the form of regression equation and comes in accord with desirable quality demands.

All the methods described in the section possess some shortcoming. Those were pointed in Doppelhofer, et al. (2000), as directions for future research. They are:

- Realization only on cross-sectional data, while panel studies have become very popular in growth literature, and sufficient base for such estimates has been developed;

- The models do not allow for possible non-linearity of some factor's influence (e.g. it has been argued that inflation has important negative effects on growth, but only for very high levels of inflation);

- Estimation is impeded by absence of some observations, and the models do not allow for unbalanced data sets.

But nonetheless these studies could be regarded as coming closer to the solving the mystery of growth. Besides, they give certain evaluation of other existing research works.

Chapter 3

OVERVIEW OF PROXIES FOR POTENTIAL GROWTH FACTORS

There have been several overviews made of the factors potentially influencing growth. In this section I am trying to bring together all possible recommendations about growth factors in general sense and those specific for transition economies. I mention their proxies commonly used in economic research, and potential channels of effect. If other is not maintained, the section is based on Frankel (1997), Levine and Renelt (1992), Sala-i-Martin (1997a), Doppelhofer, et al. (2000), Dowrick (1996) and Aron (1997).

1. General sources of economic growth.

All the factors that possibly affect long-term growth[2] are either accumulation of factors of production (physical and human capital investment), or act through investment intermediary, or improve efficiency of using resources (thus being indicators of social or organizational capital or capital accumulation), or refer to technological progress.

I mention factors together with their commonly employed proxies and special proxies with their claimants. I touch as well the specifics of calculating those proxies and channels of the factors' influence on the long-run growth. So here are, actually, variables to employ (given in brackets), grouped in classes, called factors.

[2] In per capita terms

Here only those variables are considered that, though general, can be applied to the transition countries framework (so variables like fraction of Confucians or OECD dummy are disregarded). The complete variable lists used by Levine and Renelt (1993) and Sala-i-Martin (1996a and b) can be found in the Appendices 1 and 2 accordingly.

Factor 1. Physical capital accumulation (investment as a share of GDP; public investment and private investment separately; equipment investment and non-equipment investment separately: Delong and Summers, 1991).

Influence upon growth. Facilitates growth as accumulation of a factor of production[3]; public investment improves institutional framework, lowering transaction costs.

Factor 2. Human capital accumulation (life expectancy; life expectancy squared; enrollment rates in primary and secondary school; teacher-student ratio; literacy rate; average years schooling; cognitive test scores).

Influence upon growth. Facilitates growth as accumulation of a factor of production.

Factor 3. Initial conditions (GDP^2 in the base year; $income^2$ in the base year; real labor productivity in the base year).

Influence upon growth. Describe the country's position relative to its production possibilities frontier (the bigger gap creates the greater potential). But give initial potential for future investment, opportunities for technology transfer

[3] In neoclassical theory, differential rates of investment across countries induce permanent differences in the level of per capita output, as opposed to its growth rate. In new theory (as in Romer's et al. articles of 1980s) higher rates of investment might induce higher rates of innovation, and thus faster economic growth.

and capabilities to exploit such opportunities. There is possible systematic relationship between the level of development and both returns to capital and the desire to save (Sarel 1994). There are possibly investment coordination problems and market failures (Gans 1995; Murphy et al. 1989), which implies the existence of multiple equilibria and low-growth traps.

Functional forms possible. Linear (with any sigh), U-shaped, cubic function.

Factor 4. Economically active population growth (labor force growth, population growth; fertility).

Influence upon growth. Regarded to slow growth through drawing off a part of investment. High fertility subtracts women from labor force.

Factor 5. Fiscal policy (government consumption expenditures as a share of GDP; growth of government consumption expenditures as a share of GDP; ratio of central-government tax revenue to GDP; ratio of central-government defense expenditure to GDP; central-government gross capital formation; ratio of central-government deficit to GDP; ratio of government educational expenditures to GDP; government consumption less defense and education share of GDP; ratio of social security tax to GDP).

Influence upon growth. Large budget deficits or heavy tax burdens might retard growth (thus, for example, government borrowing can crowd out private capital accumulation) etc.

Factor 6. Financial intermediation (ratio of domestic credit plus stock marker capitalization to GDP, accounting standards in the country).

Influence upon growth. Promotes saving and investment by reducing the need for self-finance; facilitates implementation highly risky projects through risk sharing; allows investing high fixed costs projects; allows for more efficient allocation of investment across competing uses.

Factor 7. Research & Development spending and spillovers (national spending for R&D; number of patents; total trade as a percent of GDP; the ratio of exports to GDP and imports to GDP separately; total tariff revenue as a percent of trade; ratio of import (export) taxes to imports (exports); measure of openness based on import penetration (residuals from regressing import share on real GDP in the base year – linear and squared); openness measure based on factor-adjusted trade: Leamer 1988; trade distortion measure based on Heckscher-Ohlin deviations: Leamer 1988; growth rate of trade; imports; exports).

Influence upon growth. Spending: creates possibilities for technological advancement; Spillover: technology transfer, expressed as availability of high-quality intermediate inputs and possibility for developing countries to replicate high-tech goods imported from developed countries; management practices transfer; spurring innovation and cost-cutting; competing away monopoly.

Factor 8. Market distortions (black-market exchange-rate premium; standard deviation of black-market exchange-rate premium; real exchange-rate distortions).

Influence upon growth. Characterize efficiency of the market operating (accounting for transaction costs etc);

Factor 9. Macroeconomic stability (the standard deviation of inflation; the standard deviation of domestic credit growth).

Influence upon growth. Any disruption of the price system causes misallocation of resources, and since leads to inefficiency; unexpected inflation discourages saving by transferring income from lenders to borrowers.

Factor 10. Monetary indicators (inflation of GDP deflator, other proxies for inflation; average growth rate of domestic credit).

Influence upon growth. High inflation tends to have great volatility, besides, measures instability itself.

Factor 11. Equality of income distribution (Gini inequality coefficient).

Influence upon growth. The electorate's demand for redistributive policies by government provokes distortionary taxation; a high concentration of wealth exasperates the credit constraints that producers face in developing countries; inequality generates social instability, having an adverse impact on growth by reducing the security of property rights.

Factor 12. Political and social conditions (frequencies of revolutions, coups, and assassinations; index of corruption; rule of law; index of bureaucratic quality; degree of capitalism; political rights index, civil liberties index).

Influence upon growth. Affect investment decisions, determine efficiency of resource allocation.

Factor 13. Primary sector production (fraction of primary products in total exports; fraction of GDP in mining);

Influence upon growth. Excessive dependence on agriculture and natural resources is likely to encourage rent seeking (thus corruption) and policy failures (thus inflation) and discourage education and external trade through the Dutch disease (Campos and Coricelli, 2000).

__Factor 14.__ Type of economic organization (index of economic freedom; index of certain sector production).

Influence upon growth. Characterizes mostly the efficiency of resource allocation in a particular society.

__Factor 15.__ Transition specific indicators (discussed beneath)

__Other Factors 16.__ (land area; urbanization rate; growth of domestic consumption expenditures; absolute latitude; index of ethnolinguistic fractionalization; fraction of population speaking foreign language; fraction of population speaking English; fraction of Jewish; ratio workers to population).

Influence upon growth. Capture scale effects, facilities of location, peculiarities of religion impact etc.

2. Specifics of growth modeling[4] in transition economies framework.

Transition process got a keen description as a period of "creative destruction": a market system is being built and in the same time the legacies of socialism are being destroyed.

[4] This section is based on Campos and Coricelli's (2000) survey of the literature on economic growth in transitional economies.

In this respect Fisher et al. (1996a) point: "a useful way to think about the current growth prospects of the transition economies is to consider them subject to two sets of forces: those arising from the transition and transformation process, and the basic neoclassical determinants of growth. The further along a country is in the transition process, the less weight on the factors that determine the transitional growth rate, and the greater the weight on the standard determinants of growth".

There has been written not very much literature on growth in transition (as compared with other categories of countries), but the literature is extending. The most important issues it touches are the relative role of initial conditions versus the role of reforms, the impact of different speed of reforms, and the role of macroeconomic variables (like inflation and the fiscal balance).

De Melo, Denizer and Gelb (1996) and De Melo and Gelb (1997) construct an index of the extent of liberalization and provide evidence that cumulative liberalization is closely related to the observed output dynamics. The index is based on the notion that it is government policies that accounts for the disparity in economic performance during transition (De Melo and Gelb, 1997). It is a weighted-average of liberalization of internal markets, of external markets and of private sector entry.

Campos and Coricelli (2000) suggest that the analysis should be carried out in terms of "transition time" rather than calendar time, using the year in which full-fledged reforms in a particular country began as starting point. It is 1991 for Central Europe and 1992 for the Former Soviet Union. Poland should be considered since 1990. The use of transition time is argued to be important because the sharp output decline was most probable in the first couple of years of transition.

Denizer (1997) points to the importance of initial conditions. He finds that initial conditions matter, as proxied by distance from Vienna and whether the country was independent before socialism.

Fisher, Sahay and Vegh (1996b) find that growth is related to initial income, the choice of exchange rates regimes, fiscal surpluses, and the (cumulative) liberalization index.

Hernandez-Cata (1997) relates the output performance to liberalization and stabilization. He finds that that the much worse performance of former Soviet Union countries, as compared to those in Central and Eastern Europe, is explained in the better part by the timing and intensity of liberalization, price stabilization and under-reporting of output.

Krueger and Ciolko (1998) and Popov (1999) show in their studies that the impact of liberalization becomes insignificant when initial conditions are accounted for.

Campos and Coricelli (2000) argue that reforms cannot be considered as simple exogenous variables, as policy choices are not exogenous decisions but rather they result from a politically constrained process affected by economic variables. Less favorable initial conditions were likely to constrain the reform process and thus push the country to follow a less radical reform path. Besides, initial conditions might have affected output performance. As a result, there could be observed a correlation between reforms and output performance even though the ultimate cause of both reforms and output performance rooted in initial conditions.

Rostowsky et al (2001) states that across transitional economies great enterprises, formerly state-owned, no matter of the way of privatization, are performing generally less efficiently than newly-created private enterprises. Therefore it may be desirable to include in the specification tested small enterprises fraction in the economy as a regressor.

One more potential source of growth is starting to exploit the productive capacities that were idle at the beginning of transition. So in the case of transition economies the case might appear even more pronounced when large capital-labor ratio would lead to a smaller GDP per capita growth. In attempt to account for these effects it seems reasonable to include growth of capital and growth of capital-labor ratio as regressors. This issue got little attention in the literature.

Summarizing, there are offered several independent variables specific to transitional economies:

1. reform indicators (internal liberalization, external liberalization, and conditions for private sector development);

2. initial conditions measures (initial distortions, macroeconomic imbalances, degree of market liberalization, natural endowment, income per capita and output growth before liberalization);

3. IMF-program dummy;

4. measures of distance from market economies, both physical and policy-determined (years spent under communism);

5. measure of small enterprise relative number in the economy;

6. proxies for idle capital.

So, a most important specific feature about studying growth in transition is the necessity to account for a number of special potentially important factors.

Chapter 4

GENERAL METHODOLOGY AND PANEL SPECIFICS

1. Specifics of applying panel models for growth equations.

Apart from factor specificity, there is another feature peculiar for studying growth in transition economies. This feature is a technical one. The maximum number of time periods today is about 12 years (or even 10 – in the sense of availability of the data), which is a rather short period, unusually short for a standard analysis in terms of long-run growth. The number of cross sections is 27 at most. This makes application of cross-section estimation impossible because of the lack of degrees of freedom. A possible way out is either cross-section model for a much broader sample of countries with a special dummy for transition, or a panel model for transition countries only.

So far, most of the growth studies conducted have been cross-sectional and not panel studies. Though, panel data models in the framework of estimating growth regressions have several important advantages as compared with cross-sectional models. Aron (1997) stresses the following two. First, in case data permit, panel data models at close to business cycle frequencies (7–10 year panels) may be used to address the vexing problems of both endogeneity and omitted variables, if sufficiently strong instruments can be found. A second important advantage of panels is that institutional change can be addressed, as well as the crucial and often frequent economic and political regime shifts, which are simply averaged out in cross-sectional studies. Besides, creating more variability, panel data alleviates multicollinearity problem and creates more efficient estimation. In the present case, although the number of cross-sections is quite different, anyway the model would produce more efficient estimators when it is panel.

Because of these reasons, the choice of the model is panel. As to choosing the appropriate estimator, there has been a lot of dispute in the literature.

Caselli et al. in the influential (1996) paper employ a 25-year panel in five-year averages, and claim to solve both the problems of correlated country effects and endogeneity, applying a GMM estimator and eliminating individual effects through first differencing. The strict assumptions made here are those of lagged values being valid instruments for income and other regressors. The shorter panels were pointed to possibly suffer biases from business cycle fluctuations. As a conclusion, the first-differenced GMM estimator was recommended for empirical growth models.

It is argued (Temple, 1996; Bond at al (2001) that the GMM estimator used by Caselli et al. does badly leading to a serious bias when the autoregressive parameters of variables employed are close to a unit root, because then lagged levels are weak instruments for current differences. Aron (1997) claims that it remains controversial whether efficient instruments are available, so that instrumental variable techniques could effectively be used to overcome the endogeneity problem.

Exploring single models with autoregressive dynamics and explanatory variables that are not strictly endogenous, Bond (2002) shows GMM to be consistent in a wide range of applications (in contrast to OLS levels and Within Groups), though lead to large biases in finite samples when the instruments available are weak.

When a study uses panel data, a common choice for macroeconomists is the fixed effects model. The reasons for that, as pointed in Judson, Owen (1996), are the following. First, if the individual effect represents omitted variables, it is highly likely that these country specific characteristics are correlated with the other regressors (which would create biased random effects estimator). Second, it is also fairly likely that a typical macro panel will contain most of the countries of interest and, thus, will be less likely to be a random sample from a much larger universe of countries.

However, there are several problems with using dummy variables to estimate individual effects. The most important one is that the use of a fixed effect estimator employing instrumental variable estimation can lead to inconsistent estimates and render invalid all but strictly exogenous variables as feasible instruments (Caselli et al, 1996). And as the framework of long-term growth modeling inherently involves double-causality problems, to eliminate inconsistency, it should necessarily engage instrumental variables estimation and is automatically subjected to the bias problem.

The conclusion here is that there are few if any estimators doing well on the short panel, especially with few cross-sections. This is the reason why in the current model the dynamic component could not be tested. Besides, to overcome the potential endogeneity problem the model employs one-period-lagged values for investment series. The choice of estimator is LSDV for unbalanced panel – within estimation (and GLS when estimating random effects in the final model), which is justified by the assumptions presented in the following section. The basic estimation formula used are presented in Appendix 3, all having basis in Baltagi (2001).

2. Methodology of General-To Specific modeling applied.

The research method proposed in this work preserves general logic of Hoover and Perez's general-to-specific modeling[5] (2001, 2002), adjusting it to panel data and using a somewhat different set of initial variables.

As inferences that are to be drawn from the resulting model would concern only those countries, which compose the data set serving the base for the final specification, these countries can be regarded population. So it would be appropriate to contract the basic model applying fixed effects estimation procedure. Besides, performing the Hausman Test at each round of specification adds computational difficulty to estimation.

It is reasonable to assume that those shocks during transition common for all the countries under consideration find their explanation in the variation of explanatory variables. So there is no need to apply the two-way error components model.

Estimation can be potentially biased at any step of the procedure because of two reasons. The first one is inappropriate model specification for the estimator chosen. The panel, available for transition countries, is neither long, nor broad. The estimators are sensitive to violating assumptions imposed on the error term. As the LS estimator is employed, it is necessary to control for autocorrelation and heteroscedasticity. The assumption is that the impact that the lagged variables have on the current value of growth is realized completely through more investment (and not resource reallocation or technological progress). So I include among regressors lagged investment and no lags of other variables, no lagged growth either. This enables escaping finite sample bias all existing estimators would have. Thus, by not including the dynamic component and paying due attention to selecting exogenous variables, we would obtain efficient estimates applying a simple estimator, for example, LSDV.

[5] Complete Hoover and Perez (2002) search algorithm is presented in the Appendix 4.

The other reason is the omitted variables bias. This problem, however, is eliminated with the great number of iterations, so the relevant variables would necessarily appear in the final outcome of the simulations. And at the later stage of performing encompassing tests this problem would exist no more. Redundant regressors at first stages do not create bias.

The assumption about the appropriate number of regressors for general regression cannot be made because of several reasons. First, and most important is the technical one. Present computing capacities do not allow computing combinations of more than four variables in the initial regression. Another one is rising multicollinearity with including more variables. At the stage of encompassing tests the number of regressors is likely to increase.

One more assumption concerns estimation with autocorrelation. It can be assumed that the serial correlation is possible of only first order. It seems reasonable to assume that any factor change would have consequences in the current ant the next period only. Its impact on the following periods is realized through other factors' changes. In principal, at the present model, the autocorrelation problem is not a grave one. The panel is short and unbalanced with observations sometimes occasionally missing in the middle of the sample.

As there are series employed in the model, which are proxies for the same factors, there is occasionally possible strong (or even perfect) multicollonearity, at the beginning of each contraction procedure current selection of independent variables is tested for multicollinearity, using Belsley procedure (as described in Greene, 1999 or Kennedy, 1998).

To check for serial correlation a variant of LM test (as in Baltagi, 2001) is applied. Test of the current regression being the valid restriction of the general regression is the F-test of joint insignificance of all the variables eliminated to get from the general regression to the current specification. Heteroscedacticity test employed is the Variant of White LM test of the regression of the residuals from the current regression on squares of all the independent variables in the current regression and their interactions.

The tests and all the estimation formula are presented in the technical Appendix 3.

The general procedure of the modeling is as follows.

1. Organize choosing combinations of L initial variables out of the set of all the potential regressos.

2. Test for possible multicollinearity and come to another combination it is detected.

3. Arrange series so as to eliminate observations missing in one of the variables, allowing for unbalanced sample.

4. Create 2 sub-samples: one including the observations for all but the last year, the other – for all but the first one. Estimate the general specification, using the full set of candidate variables and applying LSDV.

5. Run a bunch of tests: LM test for first-order serial correlation and compare it with the critical value; heteroscedasticity test; and F-test of the hypothesis that the current specification is a valid restriction of the general specification. Record the results of the tests.

6. If there are variables insignificant at more than 5 percent level, eliminate the variable with the lowest *t*-statistic and go to 3. The statistic is calculated using White's heteroscedasticity corrected standard errors.

7. If the current specification with all the variables significant at 5 percent level can pass all the tests as in 3, record it as the terminal specification and go to 11.

8. If any of the tests fails, return to the last specification, for which all the tests are passed.

9. Eliminate the variable with the lowest t-statistic, run the tests.

10. If the current specification fails any one of the tests, replace the last variable eliminated, and re-estimate the current specification eliminating the variable with the next lowest insignificant t-statistic.

11. If the current specification passes all tests, it is the terminal specification.

12. The process of variable elimination ends when a current specification passes the tests and either has all variables significant or cannot eliminate any remaining insignificant variable without failing one of the tests. Record the resulting terminal specification.

13. If there is any combination of explanatory variables different from those considered, choose it and reopen the search path.

14. Once all search paths have ended in a terminal specification, the final specification is chosen through a sequence of encompassing tests. Form the non-redundant joint model from each of the different terminal specifications; take all candidate specifications and perform the F-test for encompassing the other specifications. If only one specification passes, it is the final specification. If more than one specification passes, the specification with the minimum Schwarz criterion is the final specification. If no model passes, reopen the search on the non-redundant joint model (including testing against the general specification) using only a single search path and take the resulting model as the final specification.

15. The final specification is the intersection of the two specifications from each sub-sample.

Schematically the process of organizing the iterative process is depicted on the block-scheme at the figure below.

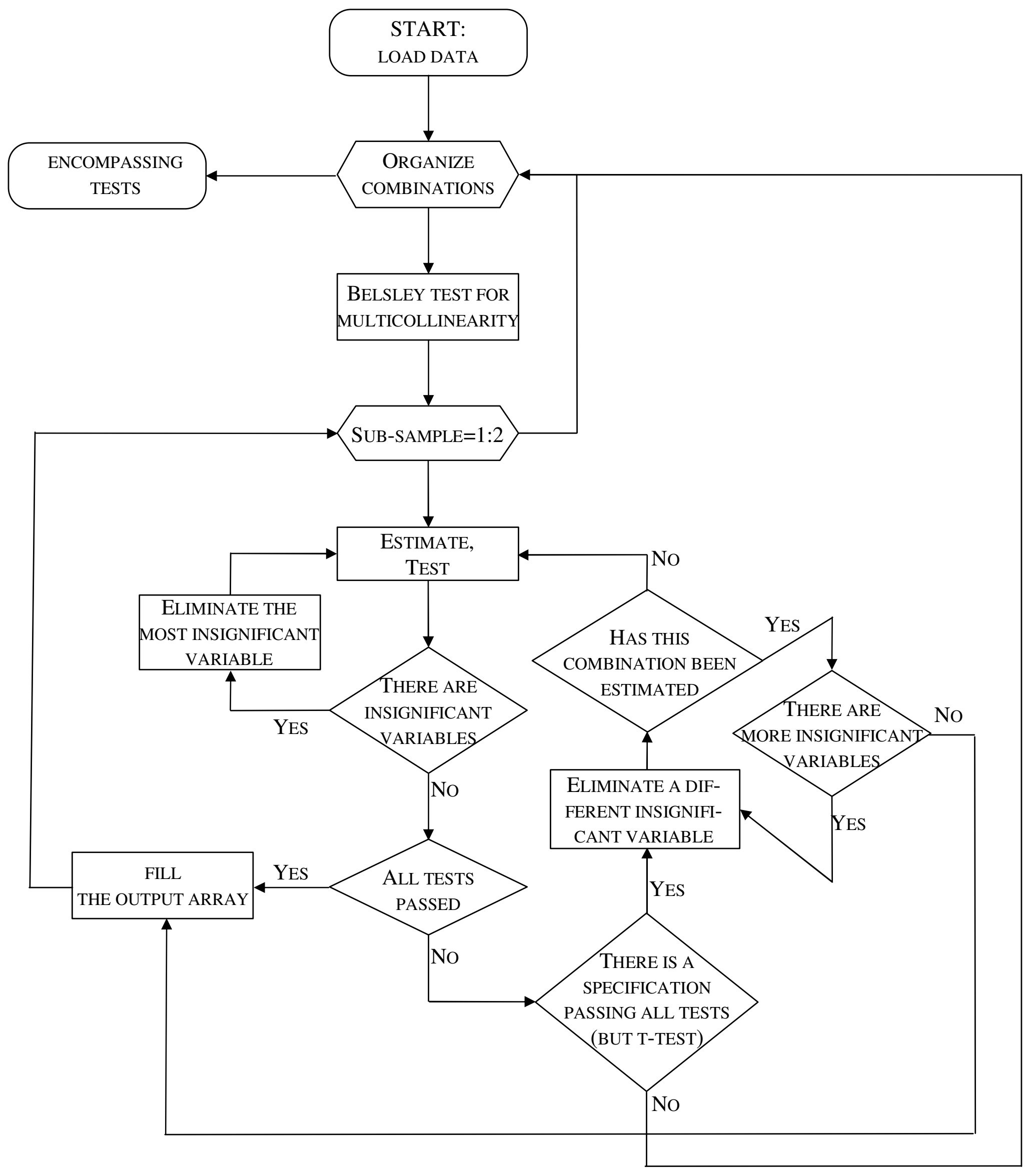

Fig.1. The Block-Scheme of the algorithm of finding robust regressors.

3. Data description.

Data is where another important feature, specific for studying growth in transition comes from. Campos and Coricelli (2000) point that there are possible limitations and inadequacies of the various time series needed among those needed. The problems with data quality and comparability are many and well documented (Bartholdy, 1997). The systemic transformation means a radical change in incentives from fulfilling plan targets to evading taxes, from over-reporting to under-reporting output. In view of this De Broeck and Koen (2000) note that, in transition, there is no "single, true real GDP series".

As was mentioned above, the data for the work comes in panel. It is yearly data for the 27 countries. They have great differences in per capita income, ranging from low income (like Ukraine or Moldova - by the World Bank classification) to high income (as Slovenia); in growth patterns, being characterized by U-shaped (CEU and Baltic countries) and L-shaped (most of CES) patterns, and so on. Nevertheless, these countries are characterized by one important thing they share in common. It is their being transition economies.

This means the countries had comparable initial conditions (expressed mostly in the similar type of economic organization), so more or less equal opportunities for economic advancement. Because of this fact they compose quite a good sample for economic and econometric endeavors. In this respect, and, besides, considering elaborating policy implications for only these economies, the assumption would be plausible that the countries chosen pepresent a whole population. Occasionally, though, because of unbalancedness of the panel, one or several countries may disappear completely from the panel.

The time span of the data is ten years for the most series: starting from the year of 1992 up to 2001. Some series are lacking observations at the beginning of the sample (such as number of patent applications to total population ratio), others – at the end (for example, index of transition progress). Some series are sparse. They are lacking occasionally data across all the sample.

Because of this fact, several series (those in principle available, but having about 50 percent of the sample), were excluded from analysis. Examples here are the R&D spendings, number of stocks traded on the stock market to GDP ratio, and some others. In many series there may be lacking data for some countries for all the years considered. Most often these countries are Yugoslavia and Bosnia and Herzegovina. Even the dependent variable series, which is the GDP per capita growth rate, is characterized by only 95.6 percent completeness because of several observations missing for Yugoslavia and Bosnia and Herzegovina, and one – for Azerbaijan of 1992.

Taking into consideration the precaution about the data, I tried to obtain series from reliable sources. Most data come from:

- World Bank's World Development Indicators;
- UNICEF's MONEE Project "Social Monitor 2002";
- EBRD's Transition Report;
- International Monetary Fund's International Finance Statistics.

Most series are used just as they come from the sources reported. Several series, however were somewhat adjusted to suit the model.

The proxy for inflation employed is year-to-year consumer price index. It is characterized by very large values in the first periods because of large inflation, sometimes measured in tens of thousand percent. Because of potential heteroscedasticity problem, the sample was truncated from above the 90th percentile of about 300 percent year-to-year. The same truncation was performed over the CPI squared series.

Proxies for investment (gross capital formation, gross fixed capital formation and foreign direct investment are taken one-period lagged. The reason for it is elimination of the endogeneity problem. Economic reason for this is the fact investment may work in the economy with some lag. Besides, previous year's investment is a good proxy for the current year's investment.

The capital stock series was calculated on the basis of PWT Methodology using WDI data for investment and economic growth (as in Parekh, 2002). The rate of depreciation was assumed 6 percent annually. Because of great range of the figures obtained, two other series were calculated. They are capital growth and capital-labor ratio growth, that when combined, would account for capital-labour ratio. This eliminates potential heteroscedasticity problem (as obtained capital per labor series have too wide range), and, besides, the latter two series are estimated more precisely.

The series of standard deviation of CPI and standard deviation of domestic credit growth were obtained by finding standard deviations (about years' means) of the IFS monthly series on CPI and domestic credit growth correspondingly.

Index of transition progress is obtained after Sachs (1994) by averaging the EBRD indices of price liberalization, trade liberalization, small-scale privatization, large-scale privatization, enterprise reform, competition policy, banking sector reform, reform of non-banking sector financial institution. The component indicators could be regarded as regressors separately, but on the rather rough grid the data are available (printed variant of Transition Report, 2002), the data are very much collinear, so it is unadvisable to have them all included in the analysis.

The complete list of the series employed in the model is presented in Appendix 5. There they come together with indication of their source, specifics of calculation, and belonging to some particular factor, topology given in Chapter 2.

Chapter 4

ESTIMATION RESULTS AND CONCLUSIONS

1. Specifics of applying panel models for growth equations.

To make the model the MatLab software was used. Organizing simulations in the number of combinations 3 out of 71 (amounting to 57155 iterations) takes the time of about half an hour. Combinations of four – about ten hours. So because of this technical constraint the estimation was provided for only 3 and 4 variables in the general regression.

After that, the results were analyzed and one more iterative procedure was done on the variables that passed the preliminary 3 and 4 initial variables combination in the general regression. Appendix 6 presents the MatLab modeling program for this final case.

In the case of 3 initial variables composing general specification there is severe problem of heteroscedasticity connected with missing variables. Still, the optimal number of regressors in the initial general regression appears to be in excess of the four (and especially three) variables employed. This is the reason of the second stage of combinations of 7 variables performed.

At such reasonable number of variables in the initial regression, encompassing tests procedure may actually lead to choosing the best specification. The problem with reopening the search path including all the variables that passed the preliminary selection stage, is the loss of degrees of freedom At the end this would be corrected by reestimating the model at the stage of performing encompassing tests. Though, reaching the specification that could encompass all the other specification is highly unlikely here.

The result of the round of simulations represents those regressors that at least once entered a "good" regression, passing all tests (and presumably having all significant variables, though not necessarily) in the both sub-samples. It includes 26 variables, describing 8 various factors on a finer grid than that classified in the above Chapter 2 (the first column corresponding to that classification number). The output is presented in the following table.

Table 1. Robust regressors in growth regressions.

Factor #	Factor name	Proxies (variables from the model)
4	Population growth related variables	1. Population growth; 2. Age dependency ratio; 3. Fertility rate.
2	Development assistance	1. Aid (% of GNI); 2. Aid (% of gross capital formation); 3. Aid (% of imports).
3	Government finance	1. Capital expenditure (% of total expenditure); 2. General government balance (per cent of GDP); 3. Military expenditure (% of GDP).
8	Current account flows	1. Imports of goods and services (annual % growth); 2. Exports of goods and services (% of GDP); 3. Exports of goods and services (annual % growth); 4. Gross foreign direct investment (% of GDP). 5. Current account balance (% of GDP); 7. External balance on goods and services (% of GDP).

1	Physical capital investment	1. Gross capital formation (% of GDP); 2. Gross fixed capital formation (% of GDP);
2	Human capital investment	1. Health expenditure, private (% of GDP); 2. Health expenditure, public (% of GDP); 3. Health expenditure, total (% of GDP); 4. Public spending on education, total (% of GDP).
15	Transition specific indicators	1. Index of Transition Progress; 2. Transition Time.
8	R&D activities	Log of Patent applications, residents
14	Structural composition of the economy	1. Industry, value added (% of GDP); 2. Agriculture, value added (% of GDP).

So out of 16 categories of variables (or factors, as denoted in chapter 2), and 14 represented by the data available, only 8 came out to produce robust regressors.

Insightfully, no one specification turned to be able to encompass all the other ones. In most of the groups of variables in the table above variables are proxies for the same factor, which is why they are likely to be collinear, and which allows choosing only one of them. Common sample showed Belsley index of multicollinearity far in excess of the adopted (in the program) figure of 30.

The reestimated model is based ultimately on the following variables: Age dependency ratio; Population growth; Aid (% of gross capital formation); General government balance (per cent of GDP); Military expenditure (% of GDP); Imports of goods and services (annual % growth); Exports of goods and services (% of GDP); Exports of goods and services (annual % growth); Gross foreign direct investment (% of GDP); Current account balance (% of GDP); Gross capital formation (% of GDP); Health expenditure, private (% of GDP); Health expenditure, public (% of GDP); Health expenditure, total (% of GDP); Public spending on education, total (% of GDP); Index of Transition Progress;

Transition Time; Industry, value added (% of GDP); Agriculture, value added (% of GDP).

At the second round of organizing combinations (where 19 of the variables in the table above were included, with reasoning coming to choosing out of similar proxies the one with the most degrees of freedom), several more variables were dropped. They were birth rate, exports, index of transition progress and transition time). Here most variables (but exports), though sharing growth's trend, have little variability and, hence, explanatory power. This is the reason why when combined with a greater number of other regressors they fail to explain anything.

The ultimate regression, obtained by performing encompassing tests looks in the following way:

$$Gr = \alpha - \underset{(0.0027)}{0.015} Aid + \underset{(0.1539)}{0.619} Gov + \underset{(0.0321)}{0.077} Exp + \underset{(0.1068)}{0.477} Inv - \underset{(0.7626)}{2.469} Pop + \underset{(0.0563)}{0.330} Educ$$

Here α is the fixed effects dummy;

Aid is Aid(% of gross capital formation);

Gov is General government balance (per cent of GDP);

Exp is Exports of goods and services (annual % growth);

Inv is Gross capital formation (% of GDP);

Pop is Population growth;

Educ is Public spending on education, total (% of GDP).

The output table of the final regression for the first sub-sample, second sub-sample and the common sample is presented in the Appendix 7. Here *Aid* stands significantly negative because of the practice of poor control he part of the IMF or other financial institutions over the loans provided, which lead just to accumulation of the foreign debt and the money was put back on private deposit accounts in the west and did not work in the economy. The issue of possible causality between growth and aid was tested by Granger Causality test performed in E-Views. Because of short time-series the lagged values were lagged two periods only (initially corresponding series were created in MatLab). The Wald test on coefficients restrictions showed that the lagged growth insignificantly affects aid. This is not a strict proof, but shows that the correlation between present aid and past economic performance is not significant.

Gov could be interpreted as the indicator of the governance quality. Higher value of *Gov* would imply more efficient allocation of resources in the economy.

Inv is physical capital accumulation. *Educ* stands as a proxy for human capital accumulation. The study shows that the classical theory of growth rooting in production function realizes in the case of transition economies, and the primary inputs are playing a significant part for economic growth.

Population growth comes with the negative sign, which is standard in growth theories. New labor force subtracts a part of the capital, marginal productivity of labor falls, so output rises less than proportionally to the labor force increase.

To conclude, in the case of transition economies the following factors affect per capita growth the most: population growth, physical and human capital accumulation, quality of the governance and (proper management of) foreign loans, all of which but the first one are to some degree policy determined. Though all the variables presented in the Table 1 above could be regarded as robust regressors in growth regressions in the case of transition economies. Because of their collinearity and proxying the same factors, only limited their number entered the ultimate specification.

The directions of further improvement in the field, which is beyond the scope of this work, could be pointed as follows. First, the data available for this work does not cover all the factors wanted. The measures of corruption, or shadow economy are presumably highly relevant. Currently there is no panel data produced for these measures long enough to be considered in this kind of the model. The inflation (especially in the first years of transition) was very much volatile, so it is a poor proxy for any monetary conditions and to measure instability as well. Potentially, some other measures of instability are to be created.

Second, when interested in the initial conditions' impact on growth one can reestimate the model applying random effect estimator and employing appropriate dummies to proxy for initial conditions.

Third, to test the results the same model could be reestimated with total factor payments as the regressant on all the independent variables but those accounting for capital and labor. For this purpose data envelopment analysis could be used.

Finally, if the methodology is to be presumed as a common method in this or other kinds of studies, recommendations should be elaborated on the appropriate number of variables to compose general specification. For this purpose it is appropriate to redo Hoover and Perez's (2002) bootstrapping simulations evaluating the probability of type 1 and type 2 errors.

BIBLIOGRAPHY

Aron, Janine. *Political, Economic and Social Institutions: A Review of Growth Evidence,* a background paper for the World Bank's 1997 World Development Report on 'The State in a Changing World,' August, 1997.

Arrelano, M. *Computing robust standard errors for within-groups estimators.* Oxford Bulletin of Economics and Statistics 49, 431-434.

Baltagi, Badi H. *Econometric Analysis of Panel Data*, John Wiley, West Sussex, 2001.

Barro, Robert J. and Jong-Wha Lee. *International Comparisons of Educational Attainment*, Journal of Monetary Economics, 32, 3, 363-394. December, 1993.

Barro, Robert J. *Determinants of Democracy*, unpublished typescript, Harvard University. 1996.

Barro, Robert J. *Economic Growth in a Cross Section of Countries*, Quarterly Journal of Economics, 106, 2 407-443, May, 199l.

Barro, Robert J., Xavier Sala-i-Martin. *Economic Growth.* McGraw Hill. 1995.

Bartholdy, Kasper, *Old and New Problems in the Estimation of National Accounts in Transition Economies*, Economics of Transition, 5 (1), pp.131-146, May, 1997.

Bhargava A., L. Franzini, W. Narendranathan, *Serial Correlation and Fixed Effects Model*, Review of Economic Studies 49, 533-549, 1982.

Bond, Stephen, Anke Hoeffler and Jonathan Temple, *GMM Estimation of Empirical Growth Models*, Journal of Economic Literature, 2001.

Bond, Stephen, *Dynamic Panel Data Models: a Guide to Micro Data Methods and practice*, CEMMAP working paper CWP09/02, April 3, 2002.

Campos, Nauro F., Fabrizio Coricelli, *Growth In Transition: What We Know, What We Don't, And What We Should*, Background Paper for the World Bank's Global Development Network, 'Explaining Growth Project', April, 2000.

Carree, Martin A., *Nearly Unbiased Estimation in Dynamic Panel Data Models with Exogenous Variables*, Tinbergen Institute Discussion Paper TI 2002-007/2, October 2001.

Caselli, Francesco, Gerard Esquivel and Fernando Lefort. '*Reopening the Convergence Debate: a New Look at Cross-Country Empirics*.' Journal of Economic Growth 1:363–369, 1996..

De Broek, Mark and Vincent Koen, *The Great Contractions in Russia, the Baltics and the Other Countries of the former Soviet Union: A view from the Supply Side*, Washington D.C.,

forthcoming IMF Working Paper, January, 2000.
De Long, J. Bradford and Lawrence H. Summers. *Equipment Investment and Economic Growth*, Quarterly Journal of Economics 106(2), 445-502, May, 1991.
De Melo, Martha and Alan Gelb, *Transition to Date: A Comparative Overview*, in Salvatore Zecchini (ed) Lessons from the Economic Transition: Central and Eastern Europe in the 1990s, Dordrecht, Kluwer Publishers, 1997.
De Melo, Martha, Cevdet Denizer, and Alan Gelb, *From Plan to Market: Patterns of Transition*, World Bank Economic Review, 10 (3) pp. 397-424, September 1996.
Denizer, Cevdet, *Stabilization, Adjustment, and Growth Prospects in Transition Economies*, Washington D.C., World Bank Policy Research Working Paper No. 1855, November 1997.
Doppelhofer, Gernot, Ronald I. Miller, and Xavier Sala-i-Martin *Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach*, NBER Working Paper No. 7750, June 2000.
Dowrick Steve *The Determinants of Long-Run Growth*, 1996.
Duffy, John and Chris Papageorgiou, *A Cross-Country Empirical Investigation of the Aggregate Production Function Specification*, Final Draft, March 2000.
Easterly William and Ross Levine, *Africa's Growth Tragedy: Policies and Ethnic Divisions*, Quarterly Journal of Economics, 112, 1203-50. November 4, 1997.
Fernandez, Carmen, Eduardo Ley, and Mark F. J. Steel *Model Uncertainty In Cross-Country Growth Regressions*, Version: January 20, 2001.
Fisher, Stanley, Ratna Sahay and Carlos Vegh, *Economies in Transition: The Beginnings of Growth*, American Economic Review Papers and Proceedings, 86 (2), May 1996a.
Fisher, Stanley, Ratna Sahay and Carlos Vegh, *Stabilization and Growth in Transition Economies: The Early Experience*, Journal of Economic Perspectives, 10 (2), pp.45-66, Spring 1996b.
Frankel, Jeffrey A., 1997 *Determinants of Long Term Growth*, Background Paper for the Morning Session of the Meeting of the Asia-Pacific Economic Cooperation Economic Advisers, Vancouver, Canada, November 20, 1997 Published as 'Why Economies Grow the Way They Do', Canadian Business Economics, Spring/Summer 1998.
Granger, Clive W. J. and Harald Uhlig, *Reasonable Extreme-Bounds Analysis*, Journal of Econometrics, 44, 159-170, 1990.
Greene, W.H. *Econometric Analysis.* Prentice Hall, New Jersey. 2000.
Hall, Robert and Charles Jones, *The Productivity of Nations*, NBER Working Paper #5812, November 1996.
Hernandez-Cata, Ernesto *Liberalization and the Behavior of*

Output during the Transition from Plan to Market, International Monetary Fund Staff Papers, 44(4), 405-29, December 1997.

Hoover, Kevin D. and Stephen J. Perez, *Evaluating Specification Search in Cross-country Growth Regressions: Extreme Bounds vs. General-to-Specific Methodologies*, Revised, 6 September 2001.

Hoover, Kevin D. and Stephen J. Perez, *Truth and Robustness in Cross-country Growth Regressions*, Journal Of Economic Literature, August 2000.

Jones, Charles I. *Introduction to Economic Growth*. New York: Norton, 1998.

Judson, Ruth A., Ann L. Owen, *Estimating Dynamic Panel Data Models: a Practical Guide for Macroeconomists*, January 1996.

Kaldor, N. *Capital Accumulation and Economic Growth*, in F.A. Lutz and D.C. Hague, eds., The Theory of Capital, New York: St. Martin's Press, 177—222, 1961.

Kennedy, Peter, *A Guide to Econometrics*, The MIT Press Cambridge, Massachusetts, 1998.

Kennedy, Peter, *Panel Data*, Kitty/Teaching/Dr Kennedy, ????

Kennedy, Peter. *A guide to Econometrics.* The MIT press Cambridge, Massachusetts. 1998.

King R., and R. Levine, *Finance and Growth: Schumpeter Might be Right*, Quarterly Journal of Economics, 1993.

King, Robert G., and Ross Levine. *Finance, Entrepreneurship, and Growth: Theory and Evidence*. Journal of Monetary Economics 32 (3): 513-42, 1993.

Kiviet, J.F*On bias, inconsistency, and inefficiency of various estimators in dynamic panel data models*, Journal of Econometrics 68, 53-78, 1995.

Knack, Stephen and Philip Keefer. *Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures*, Economics and Politics, 1995'.

Krueger, Gary and Marek Ciolko, *A Note on Initial Conditions and Liberalization during Transition*, Journal of Comparative Economics, December 1998, 26(4), pp. 718-734.

Krugman, Paul, *The Myth of Asia's Miracle*, Foreign Affairs 73, no. 6, Nov./Dec. 1994, 62-78.

Leamer, Edward E. *Sensitivity Analysis Would Help*, American Economic Review 75(3), 308-313, 1985.

Leamer, Edward E., *Measures of Openness* in R. Baldwin, ed., Trade Policy Issues and Empirical Analysis, National Bureau of Economic Research Conference Report Series, Chicago: University of Chicago Press, 1988, pp. 147-200.

Leamer, Edward E., *Measures of Openness*, in R. Baldwin, ed., *Trade Policy Issues and Economic Analysis*, National Bureau of Economic Research Conference Report Series, Chicago: University of Chicago Press, 1988, pp. 147-200.

Levine, Ross, and David Renelt, A Sensitivity Analysis of Cross-Country Growth Regressions,

American Economic Review 82, no. 4, 942-963, 1992.
Mankiw, N. Gregory, David Romer and David N. Weil. *A Contribution to the Empirics of Economic Growth. Quarterly,* Journal of Economics, 107, May, 407-437, 1992.
North, Douglass. Institutions, Institutional Change and Economic Performance. 1990.
Pack, Howard, and John Page, Accumulation, Exports, and Growth in the High-Performing Asian Economies, Carnegie-Rochester Series on Public Policy 40, 199-236, 1994.
Parekh S. Prajesh, *Determination of Capital Stock: An Analysis of Assumptions,* 5/20/02
Romer Paul M., *Human Capital and Growth*, presented at the Carnegie-Rochester Conference on Economic Policy, Rochester, NY, April 1989.
Romer, Paul M. *Capital, Labor, and Productivity.* Brookings Papers on Economic Activity, special issue, 337-420, 1990.
Sachs, Jeffrey and Andrew Warner, *Natural Resource.* Abundance and Economic Growth, mimeo HIID, 1996.
Sachs, Jeffrey D., Andrew M. Warner, *Fundamental Sources of Long-Run Growth.* The American Economic Review, Volume 87, Issue 2, Papers and Proceedings of the Hundred and Fourth Meeting of the American Economic Association, 184-188, May 1997.
Sala-i-Martin, Xavier, I Just Ran 2 Million Regressions, American Economic Review, May 1997.
Sala-i-Martin, Xavier, *I Just Ran Four Million Regressions*, National Bureau of Economic Research Working Paper Series, Working Paper # 6252, 1997.
Saltz, Ira S., and Richard J. Cebula, *An Empirical Investigation Into The Cause Of Economic Growth In The Third World Using Full Information Maximum Likelihood estimators*. The Indian Economic Journal, 1992.
Temple, Jonathan. *New Growth Evidence.* Mimeo, Hertford College, Oxford, 1996..

A

Appendix 1

LEVINE AND RENELT'S (1992) LIST OF VARIABLES

Definition	Source
Average inflation of GDP deflator	WBNA
Black market exchange rate premium	Picks Currency Yearbook; World Bank Updates
Central-government gross capital formation	IMFGFS
Dummy for OECD	
Dummy for OPEC Countries	
Dummy for outward orientation	Syrquin and Chenery (1988)
Dummy for socialist economy	Barro (1991)
Dummy variable for Latin American countries	
Dummy variable for mixed government	Barro (1991)
Dummy variable for sub-Saharan African countries	
Export share of GDP	WBNA
Government consumption less defense and education share at GDP	GOV-DEE-EDE
Government consumption share of GDP	WBNA
Growth of export share of GDP	WBNA
Growth of exports	WBNA
Growth of import share	WBNA
Growth of imports	WBNA
Growth of population	WBSI
Growth of real per capita GDP	Summers-Heston data set
Growth of real per capita GDP	WBNA
Growth of the share of government consumption (GOV)	WBNA
Growth rate of domestic credit	IMFIFS
Growth rate of government consumption expenditures	
Import share of GDP (Source: World Bank National Accounts)	WBNA
Index of civil liberties	Barro(1991)
Investment share of gross domestic product (Source: World Bank National Accounts}	WBNA

Land area (in thousands of square kilometers)	WBSI
Literacy rate in 1960	WBSI
Measure of openness based on import penetration	residuals of regression of IMP on RGDP60, RGDP602, AREA, and POP
Measure of overall trade openness	Learner, 1988
Measure of overall trade, intervention	Learner, 1988
Number of revolution and coups per year	Barro (1991)
Population in 1970	Summers and Heston (1988)
Primary enrollment rate 1960	Barro (1991)
Primary enrollment rate 1970	Barro (1991)
Ratio of central government tax revenue to GDP	IMFIFS
Ratio of central-government corporate-income-tax revenue to GDP	IMFGFS
Ratio of central-government defense expenditure to GDP	IMFGFS
Ratio of central-government deficit to GDP	IMFGFS
Ratio of central-government export-tax revenue to exports	IMFGFS
Ratio of government educational expenditures to GDP	IMFGFS
Ratio of import taxes to imports	IMFGFS
Ratio of social-security tax revenue to GDP	IMFGFS
Ratio of total government expenditure to GDP	IMFIFS
Ratio of total trade (exports + imports) to GDP	WBNA
Real exchange-rate distortion	Dollar (1991)
Real GDP per capita in 19xx	Summers and Heston(1988)
Real government capital formation	Barro (1991)
Real government consumption share of GDP	Summers-Heston data set
Real investment share of GDP	Summers-Heston data set
RERD for Summers-Heston benchmark countries	
Secondary enrollment rate 1960	Barro(1991)
Secondary enrollment rate 1970	Barro(1991)
Share of central-government individual income tax revenue to GDP	IMFGFS
Share of real government consumption expenditures minus defense and education share at expenditures	Barro (1991)
Standard Deviation of BMP	
Standard deviation of GDC (growth of domestic	IMFIFS

credit)	
Standard deviation of PI (inflation)	WBNA

WBNA: World Bank National Accounts

WBSI: World Bank Social Indicators

IMFIFS: International Monetary Fund, International Finance Statistics

IMFGFS: International Monetary Fund, Government Finance Statistics Yearbook

Appendix 2

SALA-I-MARTINE'S (1997A, 1997B) VARIABLE SET

Definition	Source or Reference	Comments
Absolute Latitude,	Barro (1996)	
Age		Average age of the population.
Area (Scale Effect).	Barro and Lee (1993)	Total area of the country.
Average Inflation Rate	Levine and Renelt (1992)	1960-90.
Average Years of Higher School	Barro and Lee (1993)	Average years of higher education of total population in 1960.
Average Years of Primary School	Barro and Lee (1993)	1960
Average Years of Schooling	Barro and Lee (1993)	1960: called "H" in definition of HUMANYL.
Average Years of Secondary School	Ban-o and Lee (1993)	1960
Black Market Premium	Ban-o and Lee (1993)	log (1+Black Market Premium).
British Colony	Barro (1996)	Dummy variable for former British colonies.
Civil Liberties	Knack and Keefer (1995)	Index.
Defense Spending Share	Barro and Lee (1993)	Fraction of GDP.
Degree of Capitalism	Hall and Jones (1996).	Index of degree in which economies favor capitalist forms of production.
Equipment Investment	Delong and Summers (1991)	
Ethnolinguistic Fractionalization	Easterly and Levine (1997)	Probability two random people in a country do not speak same language.
Exchange Rate Distortions	Barro and Lee (1993)	[Actual Source: Levine and Renelt(1992)]
Fraction of Buddhist	Barro (1996)	
Fraction of Catholic	Barro (1996)	
Fraction of Confucius	Barro (1996)	
Fraction of GDP in Mining	Hall and Jones (1996)	

Fraction of Hindu	Barro (1996)	
Fraction of Jewish	Barro (1996)	
Fraction of Muslim	Barro (1996)	
Fraction of Population Able to Speak a Foreign Language		
Fraction of Population Able to Speak English	Hall and Jones (1996)	
Fraction of Protestant	Barro (1996)	
Free Trade Openness	Barro and Lee (1993)	
French Colony	Barro (1996)	Dummy variable for former French colonies.
Government. Education Spending Share	Barro and Lee (1993)	Fraction of GDP.
Growth of Domestic Credit	Levine and Renelt (1992)	1960-90.
Growth Rate of GDP per capita	Summers/Heston Data; Penn World Tables	1960-90; the dependent variable.
Growth Rate of Population	Barro and Lee (1993)	1960-90.
H*log(GDP60)	Barro and Lee (1993)	Product of average years of schooling and log of GDP per capita in 1960.
Higher Education. Enrollment	Barro and Lee (1993)	1960.
Index of Democracy	Knack and Keefer (1995)	1965: qualitative index of democratic freedom.
Latin American Dummy		Dummy for Latin American countries.
Life Expectancy	Ban-o and Lee (1993)	1960.
Liquid Liabilities to GDP	King and Levine (1993)	Ratio of liquid liabilities to GDP (a measure of financial development).
log(GDP per capita 1960).	Barro and Lee (1993)	log(Summers-Heston GDP per capita in I960).
Non-Equipment Investment	Delong and Summers (1991)	
Number of Years Open Economy	Sachs and Warner (1996)	Index.
Outward Orientation	Levine and Renelt (1992)	
Political Assassinations	Barro and Lee (1993)	Number of political assassinations.

Political Instability	Knack and Keefer (1995)	
Political Rights	Barro (1996)	
Primary Exports	Sachs and Warner (1 996)	Fraction of primary exports in total exports in 1970.
Primary School Enrollment	Barro and Lee (1993)	1960.
Public Consumption Share	Ban-o and Lee (1993)	Public consumption minus education and defense as traction of GDP.
Public Investment Share	Barro and Lee (1993)	Fraction of GDP
Ratio Workers to Population	Barro and Lee (1993)	
Revolutions and Coups	Barro and Lee (1993)	Number of military coups and revolutions.
Rule of Law	Barro (1996)	
Secondary School Enrollment	Barro and Lee (1993)	1960.
Size Labor Force (Scale Effect).	Barro and Lee (1993)	
Spanish Colony	Barro(1996)	Dummy variable for former Spanish colonies.
Standard Deviation of Domestic Credit	King and Levine (1993)	1960-90.
Standard Deviation of Inflation	Levine and Renelt (1992)	1960-90.
Standard Deviation of the Black Market Premium	Levine & Renelt (1992)	1960-89.
Sub-Sahara African Dummy		Dummy for sub-Sahara African countries.
Tariff Restrictions	Barro and Lee (1993)	Degree of tariff barriers.
Terms of Trade Growth	Barro and Lee (1993).	1960-90.
Urbanization Rate	Barro and Lee (1993)	Fraction of population living in cities.
War Dummy	Barro and Lee (1993)	Dummy for countries that have been involved in war any time between 1960 and 1990.

Appendix 3

TECHNICAL PAGE

The basic model looks in the following way:

$$y_{it} = \alpha + x_{it} * \beta + \mu_i + \upsilon_{it} \tag{4}$$

y – the dependent variable (GDP per capita growth);

x – set of independent variables;

i – country index;

t – time index;

a – intercept;

β – vector of slope coefficients;

μ – country specific effect, $\sum \mu_i = 0$;

υ – error term.

The identical forms to present the model are (in matrix form):

$$y = \alpha + x\beta + z_{\mu} * \mu + \upsilon \tag{5}$$

$$y = Z\delta + Z_{\mu} * \mu + \upsilon \tag{6}$$

Here $Z_{\mu} = I_N \otimes i_t$;

Matrix $P = Z_{\mu} * (Z_{\mu}' * Z_{\mu})^{-1} * Z'_{\mu}$ averages observations for each cross section across time;

Matrix $Q = I_{NT} - P$ obtains deviations from individual means.

Within estimator (OLS) for unbalanced panel:

$$\tilde{\beta} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{Y} \tag{7}$$

$$\tilde{\alpha} = (\tilde{Y} - \tilde{X}\tilde{\beta}) \tag{8}$$

Here $\tilde{X} = QX$, $\tilde{Y} = QY$, $Q = diag(E_{T_i})$, $E_{T_j} = I_{T_j} - \bar{J}_{T_j}$,

$\bar{J}_{T_j} = \frac{J_{T_j}}{T_j}$, $T_j = \sum_{i=1}^{j} n_i$

GLS estimator for unbalanced panel:

$$\hat{\delta}_{GLS} = (Z^{*'}Z^{*})^{-1}Z^{*'}Y^{*} \quad (9)$$

Here $Z^* = \sigma_v \Omega^{-1/2} Z$, $Y^* = \sigma_v \Omega^{-1/2} Y$,

$\delta_v \Omega^{-1/2} = diag(E_{T_j}) + diag\left[(\delta_v / w_i)\bar{J}_{T_i}\right]$, $w_i^2 = T_i \delta_\mu^2 + \delta_v^2$

LM test for first-order serial correlation. $H_0 : \rho = 0$ **,**

Where $\rho = corr(\upsilon_{it-1}, \upsilon_{it})$.

$$LM = \sqrt{NT^2 / (T-1)}\left(\tilde{\upsilon}'\tilde{\upsilon}_{-1} / \tilde{\upsilon}'\tilde{\upsilon}\right)^2 \sim N(0,1) \quad (10)$$

F-test for validity of restrictions: *H_0*: all eliminated regressors are simultaneously insignificant.

$$F_0 = \frac{(RRSS - UKSS)/(N-1)}{URSS/(NT - K - N)} \sim F_{N-1, NT-N-K} \quad (11)$$

Hausman's specification test

$$m_1 = \hat{q}_1 \left[\mathrm{var}(\hat{q}_1)\right]^{-1} \hat{q}_1 \sim \chi_1^2 \quad (12)$$

here $\hat{q}_1 = \beta_{GLS} - \tilde{\beta}_{within}$,

var $(\hat{q}_1) = \mathrm{var}(\tilde{\beta}_{within}) - \mathrm{var}(\hat{\beta}_{GLS})$.

Random Effects model:

$$\hat{\beta}_{GLS} = \left[(X'QX / \sigma_v^2) + X'(P - \bar{J}_{NT})X / \sigma_1^2\right]^{-1} * \left[(X'QY / \sigma_v^2) + X'(P - \bar{J}_{NT})y / \sigma_1^2\right] \quad (13)$$

here $\hat{\sigma}_1^2 = T * \sum_{i=1}^{N} \bar{u}_i^2 / N$

$$\hat{\sigma}_{\nu}^{2} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(u_{it} - \overline{u}_{i})^{2}}{N(T-1)}$$

Appendix 4

HOOVER AND PEREZ'S (2001) GSM ALGORITHM

The general-to-specific search algorithm used in the simulations looks in the following way.

1. The data run are generated according to the simulated equation setup with either 0, 3, 7, or 14 true variables included. *Candidate* variables include a constant and all variables tested. A *replication* is creation of a simulated dependent variable using one of the simulated models and one draw from the bootstrapped random errors. *Nominal size* governs the conventional critical values used in all of the tests employed in the search: it is 5 percent.

2. Two sub-samples are created: one is the first 90% of the data set (ordered using the dependent variable) the other is the last 90% of the data set. Independent searches are run on the two sub-samples.

3. A *general specification* is estimated on a replication using a full set of candidate variables.

4. Five search paths are examined. Each path begins with the removal of one of the candidate variables with the five lowest t-statistics in the current general specification. All *t*-statistics are calculated using White's heteroscedasticity-corrected standard errors. The first search begins by re-estimating the regression. This re-estimated regression becomes the *current specification*. The search continues until it reaches a *terminal specification*

5. The current specification is estimated and all searchable variables are ranked according to their t-statistic. The searchable variable with the lowest *t*-statistic is removed.

6. Each current specification is subjected to the following tests:

   i. Breusch-Pagan test for heteroscedasticity

ii. subsample stability test: an F-test for the equality of the variances of the first half versus the second half of the sample in which the data are ordered according to the value of the dependent variable (which is analogous to a Chow test in a time-series context). This test compares the regressions over each subsample to the regression over the full sample. If the degrees of freedom do not permit splitting the sample into equal subsamples, the test is replaced by one that compares a regression over the first k + (n-k)/2 observations to the one over the full sample.

iii. An F-test of the hypothesis that the current specification is a valid restriction of the current general specification

7. The number of tests failed is recorded and the new specification becomes the current specification. Return to 3 until all remaining variables have a significant *t*-statistic.

8. If all variables are significant, and all of the tests in the test battery are passed, the current specification is the terminal specification and go to 10. If any of the tests fails return to the last specification for which all the tests are passed and go to 12.

9. The variable with the lowest t-statistic is eliminated. The resulting current specification is then subjected to the battery of tests.

i. If the current specification fails any one of these tests, the last variable eliminated is replaced, and the current specification is re-estimated eliminating the variable with the next lowest insignificant t-statistic.

ii. If the current specification passes all tests, re-estimate and return to 12.

iii. The process of variable elimination ends when a current specification passes the tests and either has all variables significant or cannot eliminate any remaining insignificant variable without failing one of the tests.

10. After a terminal specification has been reached, it is recorded and the next search path is tried until all have been searched.

11. Once all search paths have ended in a terminal specification, the final specification is chosen through a sequence of encompassing tests. We form the non-redundant joint model from each of the different terminal specifications; take all candidate specifications and perform the F-test for encompassing the other specifications. If only one specification passes, it is the final specification. If more than one specification passes, the specification with the minimum Schwarz criterion is the final specification. If no model passes, reopen the search on the non-redundant joint model (including testing against the general specification) using only a single search path and take the resulting model as the final specification.

12. The final specification is the intersection of the two specifications from each sub-sample.

Appendix 5

VARIABLES, EMPLOYED IN THE ANALYSIS.

# in model	Factor #	Name	Source, Footnote
Dependent variable		GDP per capita growth (annual %)	WDI
1	16	Age dependency ratio (dependents to working-age population)	WDI
2	14	Agriculture, value added (% of GDP)	WDI
3	5	Aid (% of GNI)	WDI
4	5	Aid (% of gross capital formation)	WDI
5	5	Aid (% of imports of goods and services)	WDI
6	4	Birth rate, crude (per 1,000 people)	WDI
7	1	Capital expenditure (% of total expenditure)	WDI
8	5	General government balance (per cent of GDP)	WDI
9	5	Central government debt, total (% of GDP)	WDI
10	6	Claims on private sector (annual growth as % of M2)	WDI
11	1	Capital growth	Own calculations
12	1	Capital-labor ratio growth	Own calculations
13	5	Current account balance (% of GDP)	WDI
14	5	Debt service (PPG and IMF only, % of exports of goods and services)	WDI
15	10	Deposit interest rate (%)	WDI
16	6	Domestic credit provided by banking sector (% of GDP)	WDI
17	6	Domestic credit to private sector (% of GDP)	WDI

18	3	Employment ratio (number of employed as per cent of 15-59 population)	MONEE
19	7	Exports of goods and services (% of GDP)	WDI
20	7	Exports of goods and services (annual % growth)	WDI
21	5	External balance on goods and services (% of GDP)	WDI
22	4	Fertility rate, total (births per woman)	WDI
	0	GDP per capita growth (annual %)	WDI
23	5	General government final consumption expenditure (% of GDP)	WDI
24	5	General government final consumption expenditure (annual % growth)	WDI
25	2	General secondary enrolments' (gross rates, percent of 15-18 population)	MONEE
26	11	GINI index	MONEE
27	1	Gross capital formation (% of GDP)	WDI, One-period-lagged value
28	1	Gross fixed capital formation (% of GDP)	WDI, One-period-lagged value
29	7	Gross foreign direct investment (% of GDP)	WDI, One-period-lagged value
30	2	Health expenditure, private (% of GDP)	WDI
31	2	Health expenditure, public (% of GDP)	WDI
32	2	Health expenditure, total (% of GDP)	WDI
33	2	Illiteracy rate, adult total (% of people ages 15 and above)	WDI
34	7	Imports of goods and services (% of GDP)	WDI
35	7	Imports of goods and services (annual % growth)	WDI
36	15	Index of Transition Progress	Own calculations

37	14	Industry, value added (% of GDP)	WDI
38	10	Inflation, consumer prices (annual %)	WDI
39	10	Inflation Squared	Obtained from the WDI's CPI inflation series
40	10	Interest rate spread (lending rate minus deposit rate)	WDI
41	10	Lending interest rate (%)	WDI
42	5	Life expectancy at birth, total (years)	WDI
43	10	Liquid liabilities (M3) as % of GDP	WDI
44	14	Manufacturing, value added (% of GDP)	WDI
45	5	Military expenditure (% of GDP)	WDI
46	11	Money and quasi money (M2) as % of GDP	WDI
47	11	Money and quasi money (M2) to gross international reserves ratio	WDI
48	11	Money and quasi money growth (annual %)	WDI
49	13	Ores and metals exports (% of merchandise exports)	WDI
50	13	Ores and metals imports (% of merchandise imports)	WDI
51	5	Overall budget balance, including grants (% of GDP)	WDI
52	8	LOG Patent applications, residents	WDI
53	16	Population density (people per sq km)	WDI
54	4	Population growth (annual %)	WDI
55	5	Public spending on education, total (% of GDP)	WDI
56	10	Quasi-liquid liabilities (% of GDP)	WDI
57	2	Real interest rate (%)	WDI
58	2	School enrollment, primary (% gross)	WDI
59	2	School enrollment, secondary (% gross)	WDI
60	2	School enrollment, tertiary (% gross)	WDI

61	16	Services, etc., value added (% of GDP)	WDI
62	9	Standard deviation of CPI	Own calculations based on the IMF's IFS monthly data
63	9	Standard Deviation of Domestic Credit Growth	Own calculations based on the IMF's IFS monthly data
64	5	Tax revenue (% of GDP)	WDI
65	7	Taxes on international trade (% of current revenue)	WDI
66	5	Total debt service (% of exports of goods and services)	WDI
67	5	Total debt service (% of GNI)	WDI
68	7	Trade (% of GDP)	WDI
69	16	Urban population (% of total)	WDI
70	16	Labor force to population ratio	Obtained from the WDI's Total Population and Total Labor Force series.
71	15	Transition Time	Calculated by the recommendations of Campos and Coricelli (2000)

Appendix 6

ESTIMATION PROGRAMS

Here is presented the program of estimation written in MatLab. Function called "program" is the main function organizing simulations and producing its output as a matrix of zeros and ones with rows corresponding to the terminal specifications in the successful combinations (passing all the tests employed). Following are the secondary sub-functions principally devoted to organizing the data, conducting estimation, and testing specifications.

```
function program
global A B BASE BETA C CC CSSR ctrl2 DATA ForMu Ftest HETtest ii iteration L LL
LMtest M Mcont mu2 muFULL muWAVE
global  N  NMEAN  NT  nu  NumHere  NumHere0  NumXs  optim1  optim2  output  output1
output2 pLM pt specification sub_samp
global  t  TT  TTcum  ttest  TTsub1  TTsub2  USSR  uWAVE  varBETA  vector  X  XMEAN  Xs
XWAVE y yMEAN yMEAN yWAVE
%-----------------------------------------------------------------------------
%loading data from C://mathlab/work
%-----------------------------------------------------------------------------
YY=load('Y.dat');   XX1=load('1.dat');   XX2=load('2.dat');   XX3=load('3.dat');
XX4=load('4.dat');
XX5=load('5.dat');   XX6=load('6.dat');   XX7=load('7.dat');   XX8=load('8.dat');
XX9=load('9.dat');
XX10=load('10.dat');          XX11=load('11.dat');          XX12=load('12.dat');
XX13=load('13.dat'); XX14=load('14.dat');
XX15=load('15.dat');          XX16=load('16.dat');          XX17=load('17.dat');
XX18=load('18.dat'); XX19=load('19.dat');
XX20=load('20.dat');          XX21=load('21.dat');          XX22=load('22.dat');
XX23=load('23.dat'); XX24=load('24.dat');
XX25=load('25.dat');          XX26=load('26.dat');          XX27=load('27.dat');
XX28=load('28.dat'); XX29=load('29.dat');
XX30=load('30.dat');          XX31=load('31.dat');          XX32=load('32.dat');
XX33=load('33.dat'); XX34=load('34.dat');
XX35=load('35.dat');          XX36=load('36.dat');          XX37=load('37.dat');
XX38=load('38.dat'); XX39=load('39.dat');
XX40=load('40.dat');          XX41=load('41.dat');          XX42=load('42.dat');
XX43=load('43.dat'); XX44=load('44.dat');
XX45=load('45.dat');          XX46=load('46.dat');          XX47=load('47.dat');
XX48=load('48.dat'); XX49=load('49.dat');
XX50=load('50.dat');          XX51=load('51.dat');          XX52=load('52.dat');
XX53=load('53.dat'); XX54=load('54.dat');
```

```
XX55=load('55.dat');         XX56=load('56.dat');         XX57=load('57.dat');
XX58=load('58.dat'); XX59=load('59.dat');
XX60=load('60.dat');         XX61=load('61.dat');         XX62=load('62.dat');
XX63=load('63.dat'); XX64=load('64.dat');
XX65=load('65.dat');         XX66=load('66.dat');         XX67=load('67.dat');
XX68=load('68.dat'); XX69=load('69.dat');
XX70=load('70.dat'); XX71=load('71.dat');                                    %
?CH?
%-----------------------------------------------------------------------
%       transforming the tables into vectors XXX(I)
%-----------------------------------------------------------------------
XX=cat(3, XX1, XX2, XX3, XX4, XX5, XX6, XX7, XX8, XX9, XX10, XX11, XX12, XX13,
XX14, XX15, XX16, XX17, XX18, XX19, XX20, ...
    XX21, XX22, XX23, XX24, XX25, XX26, XX27, XX28, XX29, XX30, XX31, XX32,
XX33, XX34, XX35, XX36, XX37, XX38, XX39, XX40, ...
     XX41, XX42, XX43, XX44, XX45, XX46, XX47, XX48, XX49, XX50, XX51, XX52,
XX53, XX54, XX55, XX56, XX57, XX58, XX59, XX60, ...
      XX61, XX62, XX63, XX64, XX65, XX66, XX67, XX68, XX69, XX70, XX71);
  clear XX1 XX2 XX3 XX4 XX5 XX6 XX7 XX8 XX9 XX10 XX11 XX12 XX13 XX14 XX15 XX16
XX17 XX18 XX19 XX20 ...
    XX21 XX22 XX23 XX24 XX25 XX26 XX27 XX28 XX29 XX30 XX31 XX32 XX33 XX34 XX35
XX36 XX37 XX38 XX39 XX40 ...
    XX41 XX42 XX43 XX44 XX45 XX46 XX47 XX48 XX49 XX50 XX51 XX52 XX53 XX54 XX55
XX56 XX57 XX58 XX59 XX60 ...
    XX61 XX62 XX63 XX64 XX65 XX66 XX67 XX68 XX69 XX70 XX71;
    %specification=0;           %|           size of the output vector
    %output=[];                 %|           output vector
    %output1=[]; output2=[];      %|              outputs of regressions for
different sub-samples
%-----------------------------------------------------------------------
%       organizing combinations
%-----------------------------------------------------------------------
Xs=[1 2 4 6 7 8 13 19 20 27 29 30 31 35 36 37 54 55 71];
NumXs=length(Xs);
for i1=2,                     %| choosing the first regressor variable
for i2=i1+1:NumXs-5,
for i3=i2+1:NumXs-4,
for i4=i3+1:NumXs-3,
for i5=i4+1:NumXs-2,
for i6=i5+1:NumXs-1,                    %| choosing the last regressor variable
?CH?
for i7=i6+1:NumXs
    ii=[i1   i2   i3   i4   i5   i6   i7]                                   %|
?CH?
    NumHere=Xs(ii);                                                          %
?CH?
    NumHere0=NumHere;
    L=length(NumHere);                    %| Number of variables composing the
general specification
%-----------------------------------------------------------------------
%       checking restrictions on multicollinearity
%       through Belsley (1980) procedure
%-----------------------------------------------------------------------
    MULT=[];
    for jj=NumHere,                                %| creating set of current
regressors
        MULT=[MULT, reshape((XX(:,:,jj))',[],1)];
    end
    k=1:L;
    S=diag(diag(ones(L)./(((MULT(:,k))'*MULT(:,k))).^0.5));
    Lambda=eig(S*(MULT'*MULT)*S);
    Gamma=sqrt(max(Lambda)/min(Lambda));    %| condition number
```

```
    iii=(Gamma>20)|(Gamma==Inf);                  %| string of "truth" if the
considered
                                              %| variables are present together
%--------------------------------------------------------------------------
%       Creating current data set
%--------------------------------------------------------------------------
    if iii==0                                     %| checking multicollinearity
restrictions
        NN=(1:size(YY,1));                  %| array of countries'numbers
        A=[reshape(YY',[],1), MULT];        %| array of current variables
        A(find((sum(A'==0))'),:)=zeros;
        C=A; CC=C;
        create_DATA;                        %| creating the array DATA
        A=[A (1:size(A,1))'];
        A(find((sum(A'==0))'),:)=[];        %| Deleting rows with empty cells
        PPP=zeros(10,27);                                                  %|
?CH?
        PPP(A(:,end))=1;
        TT=sum(PPP);
        B=A;                               %| Solid raw data double array
        TT(TT==0)=[];                        %| number of observations for each
country
        N=length(TT);                             %| Number of countries under
consideration
%--------------------------------------------------------------------------
%       organizing calculation cycles
%--------------------------------------------------------------------------
        sub_samp=1;
        while sub_samp==1|sub_samp==2,        %| Coming to itterations on the
first sub-sample
            arrange_data;                   %| Arranging data
            ctrl2=0; ctrl1=0;               %| control variables
            while ctrl2==0,                   %| Organizing primary elimination
path
%--------------------------------------------------------------------------
%       calculatrions
%--------------------------------------------------------------------------
                calculus;               %| basic estimation
                LM_test;                %| LM test for autocorrelation
                HET_test;               %| Heteroscedasticity test
                F_test;                   %| F-test for joint insignificance of
the regressors
                t_test                  %| t-test
%--------------------------------------------------------------------------
%       Primary eliminating the most insignificant variable
%       and rearanging initial data
%--------------------------------------------------------------------------
                if max(pt)>0.05              %| do not come if all X's are
significant
                    if LL>1                  %| or there is nothing more to
eliminate
                        iteration=iteration+1;
                        DATA(1,:,iteration)=DATA(1,:,(iteration-1));
                        if ctrl1==1
                            check_previous_iterations(iteration-1);
                            if max(vector)>0.05          %whether there are
combinations unchecked
                                rearange_A
                            else
                                DATA(:,:,iteration)=[];
                                iteration=iteration-1;
                                ctrl2=1;     %if there are no combinations
unchecked exit the cycle
```

```
                        end
                    else
                        DATA(1,find(DATA(2,:,iteration-
1)==max(DATA(2,:,iteration-1))),iteration)=0;
                        rearange_A
                    end
                else
                    ctrl2=1;     %| if there are no variables to eliminate
exit the cycle
                end
            else
                ctrl2=1;         %| if all the variables passed the t-test
exit the cycle
            end
        end                     %| end ctrl2
%--------------------------------------------------------------------------
%       Checking whether the tests are passes after contraction
%       and filling the output matrices
%--------------------------------------------------------------------------
        ctrl1=1;
        if (ttest==ones)|(LL==1)         % if the former is not true - go
another iteration
                                         %  if  the  latter  -  othe  current
combination has no output
            if   (LMtest==1   &   HETtest==1   &   Ftest==1);   %   checking
specification's goodness of fit
                fill_output
            else                        % qq is the smallest specification,
passing all tests
                qq=max(find(sum(reshape(DATA(3,1:3,:),3,[]))==3));
                if qq>0
                    iteration=iteration+1;
                    DATA(1,:,iteration)=DATA(1,:,qq);
                    check_previous_iterations(qq)
                    if max(vector)>0.05&length(find(DATA(1,:,iteration)))>1
                        rearange_A
                    else
                        fill_output
                    end
                else
                    ctrl2=1; sub_samp=3;
                end
            end
        else
            ctrl2=1; sub_samp=3;
        end                 %end ttest==ones...
    end         % end choosing sub-samples
  end           % end checking multicollinearity restrictions
end
end
end
end
end
end
end
save Cap6.mat output


%--------------------------------------------------------------------------
%          creating the DATA array
%--------------------------------------------------------------------------
function create_DATA
%           DATA is the current itteration data in which
```

```
%           the first line is the X's chosen,
%           the second - p-values of t-statistics,
%           the third - 1: result of the LM test (1=pass),
%                       2: result of the White test,
%                       3: result of the F-test
%                       4: # of x previously eliminated
global L LL DATA iteration NumXs NumHere NumHere0 ii;
NumHere=NumHere0; LL=L;
iteration=1;
DATA=zeros(3,NumXs);
DATA(1,ii)=1;

%---------------------------------------------------------------------------
%        Arranging data
%---------------------------------------------------------------------------
function arrange_data
global A B TT sub_samp N L LL M y X yWAVE XWAVE NT yMEAN XMEAN TTcum iteration
sub_samp Mcont TTsub1 TTsub2 C CC ForMu BASE mu2;
if iteration==1
    C=CC; A=B;
    if sub_samp==1
        TTsub1=cumsum(TT);
        TTsub2=cumsum(TT)-TT(1)+1;
        TTsub1(find(TT<7))=[];
        TTsub2(find(TT<7))=[];
        TT(find(TT>6))=TT(find(TT>6))-1;                      %vector of years
considered for all countries
        TTcum=cumsum(TT);
        C(A(TTsub1,end),:)=zeros;
        A(TTsub1,:)=[];
    else
        C(A(TTsub2,end),:)=zeros;
        A(TTsub2,:)=[];
    end
    BASE=permute((reshape(C',LL+1,[],27)),[2 1 3]);                          %
?CH?
    Mbase=sum(BASE)./sum(logical(BASE));
    M=(reshape(Mbase,LL+1,[]))';
    ForMu=(isfinite(sum(M')))';
    ForMu=ForMu.*((1:size(ForMu,1))');
    mu2=(reshape((permute(logical(BASE),[2,1,3])),LL+1,[]))';
    mu2(:,2:end)=[];
    ForMu(ForMu==0)=[];
    M(find((sum(isnan(M')))'),:)=[];
    Mbase=repmat(Mbase,[10,1,1]);
    MeanBase=(reshape((permute(Mbase,[2,1,3])),LL+1,[]))';
    MeanBase(find((sum(C'==0))'),:)=[];
    C=A(:,1:LL+1)-MeanBase;
    y=A(:,1);
    yMEAN=M(:,1);
    yWAVE=C(:,1);
    X=A(:,2:LL+1);
    XMEAN=M(:,2:end);
    XWAVE=C(:,2:end);
    Mcont=M;
    NT=length(y);                                 %| Degrees of freedom of
the data set
end

%---------------------------------------------------------------------------
%       basic estimation
%---------------------------------------------------------------------------
function calculus
```

```
global BETA XWAVE yWAVE muWAVE uWAVE NT varBETA yMEAN XMEAN LL USSR CSSR N
muFULL nu y X L TT C ForMu BASE mu2;
BETA=inv(XWAVE'*XWAVE)*XWAVE'*yWAVE;
mu=yMEAN-XMEAN*BETA;
uWAVE=yWAVE-XWAVE*BETA;
j=1:NT;
varBETA=inv(XWAVE'*XWAVE)*XWAVE'*diag((uWAVE(j)).^2)*XWAVE*inv(XWAVE'*XWAVE);
if LL==L
    USSR=uWAVE'*uWAVE;
else
    CSSR=uWAVE'*uWAVE;
end
mu1(1:27)=NaN;
mu1(ForMu)=mu;
muFULL=reshape(kron(mu1,ones(10,1)),[],1);
muFULL(find(mu2==0))=[];
nu=y-X*BETA-muFULL;                          %residuals for the nontransformed
data

%------------------------------------------------------------------------
%       LM test for autocorrelation
%------------------------------------------------------------------------
function LM_test                             %modified LM with possible cross-
section correlations eliminated
global nu TT N NT pLM LMtest DATA iteration TTcum;
nu1=nu; nu2=nu; a1=0; a2=0;
nu1(TTcum-TT(1)+1)=[];
nu2(TTcum)=[];
LM=(NT-N)*sqrt(1/(NT-N-1))*((nu2'*nu1)/(nu2'*nu2))^2;
pLM=2*(1-cdf('Normal',LM,0,1));
if pLM<0.1                                           % if so reject H0 on no
autocorrelation
    LMtest=1;
    DATA(3,1,iteration)=1;                   % putting the result into the
current selection data
else
    LMtest=0;
end

%------------------------------------------------------------------------
%       Heteroscedasticity test
%------------------------------------------------------------------------
function HET_test
global NT LL X nu HETtest iteration DATA;
if LL>1
    Comb=combntns(1:LL,2);
    Comb1=Comb(:,1);
    Comb2=Comb(:,2);
    XHET1=[]; XHET2=[];
    XHET1(:,1:length(Comb1))=X(:,Comb1(1:length(Comb1)));
    XHET2(:,1:length(Comb2))=X(:,Comb2(1:length(Comb2)));
    XHET=XHET1.*XHET2;
    XHET=[XHET, X.*X];
else
    XHET=X.*X;
end
BETAmu=inv(XHET'*XHET)*XHET'*(nu.^2);
EPSILONmu=nu.*nu-XHET*BETAmu;
R2=1-(EPSILONmu'*EPSILONmu)/((nu.^2)'*(nu.^2));
pR2=1-cdf('chi2',NT*R2,size(XHET,2));
if pR2>0.05                                          %if so reject H0 on no
heteroscedasticity
    HETtest=1;
```

```
    DATA(3,2,iteration)=1;             %| putting the result into the current
selection data
else
    HETtest=0;
    DATA(3,2,iteration)=0;
end

%-------------------------------------------------------------------------
%       F-test for joint insignificance of the regressors
%-------------------------------------------------------------------------
function F_test
global CSSR USSR NT L LL Ftest DATA iteration;
if LL==L
    Ftest=1;
else
    F=((CSSR-USSR)/(L-LL))/(USSR/(NT-L));
    pF=1-cdf('F',F,(L-LL),(NT-L));
    if pF>0.05                                       %if so reject H0 on valid
restriction
        Ftest=1;
    else
        Ftest=0;
    end
end
if Ftest==1
    DATA(3,3,iteration)=1; % putting the result into the current selection data
end

%-------------------------------------------------------------------------
%       t-test
%-------------------------------------------------------------------------
function t_test
global L BETA varBETA pt t NT LL ttest NumXs DATA iteration ii;
t=(BETA./sqrt(diag(varBETA)))';
pt=2*(1-cdf('t',abs(t),(NT-LL)));
ttest=zeros(1,LL);
ttest(find(pt<0.05))=1;
DATA(2,find(DATA(1,:,iteration)),iteration)=pt;

%-------------------------------------------------------------------------
%       composing the ultimate output
%-------------------------------------------------------------------------
function fill_output
global sub_samp optim1 DATA optim2 NumXs output output1 output2 specification
iteration ctrl2;
global L LL DATA iteration NumXs NumHere NumHere0 ii;
if sub_samp==1                         %|
    optim1=DATA(1,:,size(DATA,3));     %|
    iteration=1;
    create_DATA;
else                                   %|
    optim2=DATA(1,:,size(DATA,3));     %|
    for q=1:NumXs,                     %|
        if optim1(q)~=optim2(q)        %|  if the second sub-sample passes
            optim(q)=0;                  %|  all tests save the output as the
intersection
        else                                 %|  of the two and come to new
iteration
            optim(q)=optim1(q);        %|
        end                            %|
    end                                %|
    specification=specification+1;
    output=[output; optim];            %|
```

```
    output1=[output1; optim1];
    output2=[output2; optim2];
end                                     %|
sub_samp=sub_samp+1; ctrl2=1;


%-------------------------------------------------------------------------
%       choosing combination not explored previously
%-------------------------------------------------------------------------
function check_previous_iterations(ss)
global DATA iteration sub_samp vector;
cv=DATA(1,:,iteration);
vector=DATA(2,:,ss);
zz=0;
vector(find(vector==max(vector)))=0;
while zz==0;
    cv(find(vector==max(vector)))=0;
    if                find(sum(reshape(DATA(1,:,1:iteration-1),[],iteration-1)-
repmat(cv,iteration-1,1)')==0)>0
        cv(find(vector==max(vector)))=1;
        if vector(find(vector==max(vector)))<0.05
            zz=1;
        end
        vector(find(vector==max(vector)))=0;
    else
        zz=1;
        DATA(1,:,iteration)=cv;
    end
end

%-------------------------------------------------------------------------
%       putting A in accordance with current restrictions
%-------------------------------------------------------------------------
function rearange_A
global DATA A iteration NumHere NumHere0 X XMEAN XWAVE C M L Mcont LL ii Xs;
NumHere=Xs(find(DATA(1,:,iteration)==1));
LL=length(NumHere);
X=A(:,2:L+1);
X(:,find(DATA(1,ii,iteration)==0))=[];
M=Mcont;
M(:,find(DATA(1,ii,iteration)==0)+1)=[];
XMEAN=M(:,2:end);
XWAVE=C(:,2:end);
XWAVE(:,find(DATA(1,ii,iteration)==0))=[];
```

Appendix 7

ESTIMATION OUTPUT

Table A7.1. Final regression estimation.

		Aid	**Gov**	**Exp**	**Inv**	**Pop**	**Educ**
1 sub-sample	**Beta**	-0.0153	0.5760	0.0825	0.4886	-2.8648	0.3623
	St.error	0.0029	0.1784	0.0325	0.1069	1.0233	0.0664
	p-value	0.0000	0.0016	0.0124	0.0000	0.0059	0.0000
2 sub-sample	**Beta**	-0.0576	0.6673	0.0816	0.5221	-2.5844	0.2872
	St.error	0.0165	0.1548	0.0314	0.1096	0.6171	0.0471
	p-value	0.0007	0.0000	0.0103	0.0000	0.0001	0.0000
Common sample	**Beta**	-0.0149	0.6189	0.0774	0.4772	-2.4692	0.3304
	St.error	0.0027	0.1539	0.0321	0.1068	0.7626	0.0563
	p-value	0.0000	0.0001	0.0173	0.0000	0.0015	0.0000

Results of Granger test:

Table A7.2. Regression output:

Dependent Variable: GROWTH
Method: Least Squares
Date: 05/26/03 Time: 14:57
Sample: 1 213
Included observations: 213
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LAG1GROWTH	-1.205276	1.069422	-1.127035	0.2610
LAG2GROWTH	-0.308420	0.624964	-0.493501	0.6222
LAG1AID	0.367901	0.281284	1.307934	0.1923
LAG2AID	-0.255040	0.179195	-1.423252	0.1562

R-squared	0.226798	Mean dependent var	8.99E-05
Adjusted R-squared	0.215699	S.D. dependent var	72.79324
S.E. of regression	64.46625	Akaike info criterion	11.18876

Sum squared resid	868582.5	Schwarz criterion	11.25188
Log likelihood	-1187.603	Durbin-Watson stat	1.775825

Table A7.3. Test results.

Wald Test:
Equation: EQ01

Test Statistic	Value	df	Probability
F-statistic	1.194193	(2, 209)	0.3050
Chi-square	2.388386	2	0.3029

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(3)	0.367901	0.281284
C(4)	-0.255040	0.179195

Restrictions are linear in coefficients.