DETERMINANTS OF US FDI AND ECONOMIC GROWTH IN SUB-SAHARAN AFRICA

NGUYEN MINH HUY DUONG

A thesis submitted in partial fulfilment of the requirements of the University of the West of England, Bristol for the degree of Doctor of Philosophy

Faculty of Business and Law, University of the West of England, Bristol

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

ADF	Augmented Dickey-Fuller
BEA	Bureau of Economic Analysis, the US
BHT	Bond, Hoeffler and Temple (2001)
BL1.3	Educational Attainment Dataset version 1.3 published by Barro and Lee (2013)
CCE	Common Correlated Effects
CCEMG	CCE Mean Group
CCEP	CCE Pooled
CEL	Caselli, Esquivel and Lefort (1996)
CIF	Costs, Insurance and Freight
CTPA	Centre for Tax Policy and Administration, OECD
DGMM	First-Differenced Generalised Method of Moments
DGMM2R	Two-step First-Differenced Generalised Method of Moments with Windmeijer-Corrected Standard Errors
FDI	Foreign Direct Investment
FE	Fixed Effects
FOB	Free on Board
GDP	Gross Domestic Product
GMM	Generalised Method of Moments
HCASM	Human Capital Augmented Solow Model
I(1)	Integrated of Order One
IFS	International Financial Statistics, IMF
IMF	The International Monetary Fund
LLC	Levin, Lin and Chu
LSDVC	Bias-Corrected Least Square Dummy Variable
MG	Mean Group
MNF(s)	Multinational Firm(s)
MRW	Mankiw, Romer and Weil (1992)
M-W	Maddala-Wu
OECD	Organisation for Economic Co-operation and Development
OECDStat	Statistics Database, OECD
OLS	Ordinary Least Squares

POLS	Pooled OLS
PWT80	Penn World Table version 8.0 published by Feenstra, Inklaar and Timmer (2013)
RE	Random Effects
RMSE	Root Mean Squared Error
SGMM	System Generalised Method of Moments
SGMM2R	Two-step System Generalised Method of Moments with Windmeijer-Corrected Standard Errors
UK	United Kingdom
UNCTAD	United Nations Conference on Trade and Development
UNCTADstat	Statistics Database, UNCTAD
US	United States of America
USD	United States Dollars
WDI	World Development Indicators, World Bank
WG	Within Groups

ACKNOWLEDGEMENT

The PhD program is a long and tough journey. In this section, I would like to take this opportunity to thank all those who have supported and helped me in the journey.

I firstly express my profound gratitude to my advisors, Jonathan Temple and Iris Mariscal, for their valuable guidance, assistance, support and encouragement throughout the program. Without them, the research could not have been possible.

I am grateful to Markus Eberhardt, Hashem Pesaran, David Roodman, Stephen Bond and Alexander Chudik for their helpful guidance and comments relevant to empirical methods in my thesis; additionally, thank Markus Eberhardt and David Roodman very much for explaining and sharing their STATA codes with me. I am also thankful to Robert Feenstra and Robert Barro for their explanation and guidance on PWT 8.0 and BL 1.3 data sets.

Special thanks go to Jon Tucker, Peter Howells, Cherif Guermat and Michail Veliziotis who are very kind to read, correct and make useful comments and suggestions throughout the thesis. Besides, I greatly appreciate the encouragement and help from Svetlana Cicmil, Paul Dowdall, Don Webber, Helen Jackson, Helen Frisby, Helen Mulligan and Samantha Watts during the last year.

I thank my examiners, Geoff Pugh and Felix Ritchie, for their constructive comments in review of my thesis.

Also, I am thankful to Department of Accounting, Economics and Finance, Faculty of Business and Law, Graduate School and Frenchay Library, UWE for providing me a stimulating study environment and allowing me to write the thesis.

Thank my fellow PhD students at UWE, especially Vasco Vendrame, Ganess Dirpal, Bolanle Karimu and Benoit Dutilleul, for helpful discussions as well as for the laughter and tear we share. Besides, I thank my friends at UWE, Vietnamese Student Association at UWE, UWE Students' Union, General Vietnamese Student Association and Embassy of Vietnam in the UK for being side by side with me during my time of study in Bristol, UK with enthusiastic help and support.

Last but not least, I would like to express my deepest gratitude to my family, grandmother, parents, parents-in-law, my wife and unborn son for their love, patience, support and encouragement, especially for the last busy years. The birth of my son expected in June is a motivation for me to complete this thesis.

Dedication

To my maternal grandfather, Nguyen Van Nhung (1920-1955), and my uncle, Duong Ngoc Sang (1953-1973), who I never met but they always live forever in my mind.

To my beloved octogenarian grandmother, Duong Thi Huong, who always supports me with her constant love.

– As long as a branch of science offers an abundance of problems,
so long is it alive – David Hilbert, 1900

ABSTRACT

This thesis consists of two separate studies as follows.

The first study uses macro panel data on US FDI in developed countries during 1982-2010 to empirically investigate the influence of host country characteristics on FDI. Differing from earlier panel data studies on FDI determinants which often impose the (standard) restrictions of the homogeneity of slope coefficients on the observed variables and the homogeneity of the factor loadings on the unobserved common factors in the empirical specification, this study uses the recently-introduced Common Correlated Effects Mean Group estimator to allow the effects of observed variables and unobserved common factors to vary across countries. In this research, the data seem to support the empirical specification allowing for slope heterogeneity across countries rather more than the standard ones imposing the restrictions of slope homogeneity. Empirical results indicate that the stock of US FDI in a given FDI recipient is likely to be significantly determined by market size, the fluctuations of the exchange rate, and risks in terms of the investment climate, corruption and the legal environment of the host country.

The second study uses an efficient two-step system GMM estimator with Windmeijer-corrected standard errors to test the human-capital augmented Solow model (HCASM). Empirical results in this study confirm conditional convergence as the HCASM predicts. However, the rate of convergence found ranges from about 0.3 to roughly 1 per cent a year, which is slower than found in previous cross-country research. The effect of the investment rate on the level of the growth path is found to be significant while that of the level of human capital is insignificant. Besides, this study finds that the HCASM seems to be unable to fully account for the contrasting growth of countries in sub-Saharan Africa and East Asia even when country-specific effects and endogeneity are taken into account. Further, the evidence indicates that the rates of technological progress between the two regions are likely to be different and this may help to explain the contrasting growth performance experienced by sub-Saharan African and East Asian countries.

STUDY I:

THE DETERMINANTS OF US OUTWARD FDI INTO DEVELOPED COUNTRIES SINCE 1982

CHAPTER I.1:
INTRODUCTION

The aim of this study is to investigate the influence of host country factors on foreign direct investment (FDI) from the United States (US) to developed countries in the Organisation for Economic Co-operation and Development (OECD) from 1982 until 2010. This study differs from previous panel data studies on FDI determinants in that it uses the recently-introduced Common Correlated Effects Mean Group (CCEMG) estimator to allow the effects of observed variables and unobserved common factors to vary across countries.

There are two major problems relating to the estimation methods in the FDI-determinant literature using aggregate panel data at the country level. Firstly, earlier studies, for example Nigh (1985) and Koechlin (1992), often use standard panel estimation methods such as Pooled OLS (POLS) or Fixed Effects (FE). In these cases, the slope parameters for the observed explanatory variables are typically constrained to be constant across recipients. This restriction can be too strong since the impact of a given factor on FDI may be different for different recipients. Given that the observed panel samples have a long time series dimension in this study, it could be more informative to allow the parameters to be heterogeneous across recipients.

Secondly, previous macro-panel studies have often controlled for unobserved common factors with the restriction that the effect of the common factors is homogeneous across FDI recipients. However, common factors are likely to be diverse. For example, they may be global events such as the recent financial crisis. Additionally, in the context of investigation of the determinants of FDI from an investing country to a cross-section of host countries, the common factors could also be related to advanced knowledge, technological expertise or superior managerial systems of the investing country’s firms. Those factors may affect FDI and thus they should be accounted for. Using aggregate country-level panel data, it may be difficult to correctly measure the common factors, but failing to control for them could lead to misleading inference. Furthermore, since common factors are diverse, it is reasonable to believe that the impact of them could be heterogeneous across different host countries.[1] This study aims to fill the gap in the literature by using the recently-developed CCEMG estimator to allow the influence of observed variables and unobserved common factors to vary across countries. To our knowledge, this estimator has not previously been applied in the context of FDI-determinant studies.[2]

There are three reasons why a panel sample of US FDI in developed OECD countries is used here. Firstly, according to statistical data from the UNCTADstat database of the United Nations, the US has been the world's largest investing country, in terms of FDI, for several decades. Secondly, developed OECD countries are the largest recipients for total global FDI. Data from the UNCTADstat database show that the stock of FDI in OECD countries accounts for approximately seventy percent of the total FDI stock of the world over the last thirty years. As for the overseas direct investment of US firms, OECD countries are also the largest destination, accounting for more than two-thirds of the total stock of outbound US FDI for the last three decades according to the US Bureau of Economic Analysis (BEA). Thirdly, annual data on variables in our model are likely to be more reliable and available

---

[1] For example, the average advantages of technology, knowledge and managerial skills of the US over Germany may be different from those of the US over Greece or Portugal. Thus, the impacts of the advantages of the technology, knowledge and managerial skills on US FDI to the host countries such as Germany, Greece and Portugal may be heterogeneous.

[2] The following key words were used in the database of the Research Papers in Economics (RePEc), the database Econlit and Google Scholar: "common correlated effects mean group", "foreign direct investment", "determinants"; "common correlated effects mean group", "foreign direct investment", "determinants", "United States"; "common correlated effects mean group", "foreign direct investment", "determinants", "United States", "OECD".

consecutively for the developed countries in comparison with developing countries.

The focus of this study is the analysis of the determinants of US FDI to OECD countries rather than a two-way study. The reasons are as follows. Firstly, one of the limitations of such a two-way study is that the parameters for a given recipient country are likely to vary across investing countries, and hence in practice, a flexible approach to a two-way study is likely to require that models are estimated separately (investing) country-by-country. By focusing on just FDI from the US to host countries, this study can examine that case in more depth and allow for slope parameter heterogeneity in the estimates. Also, the focus on US FDI to a cross-section of countries enables us to control for factors such as the (average) relative knowledge or technology advantages of the investing country's firms as unobserved common factors in the analysis.

The remainder of this study is as follows. Chapter two presents the literature review. Chapter three discusses the model, data and empirical methods used in this study. Chapter four reports empirical results and, finally, chapter five provides the conclusion of the research.

CHAPTER I.2:
LITERATURE REVIEW

There are a number of theories of FDI. For example, the database of the Research Papers in Economics (RePEc) lists over eight thousand references for foreign direct investment. Therefore, this literature review can be only selective. Before the 1960s, most theories such as Iversen (1935) and Markowitz (1959) explained overseas investment based on the assumption of perfect markets. However, Hymer (1976) and Kindleberger (1969) argued that in a perfectly competitive market, all firms compete equally and have no advantages over each other, so that FDI has no reason to exist. In his doctoral thesis of 1960, later published in 1976, Hymer showed that firms operating in foreign markets often face a variety of disadvantages compared to indigenous firms, for example language differences or lack of customer tastes. Faced with these disadvantages, for a firm to engage in investment in foreign markets, it must possess specific ownership advantages such as knowledge or technology to balance the disadvantages of operation in a foreign country. Specific-ownership advantage is a source of market power to help a firm to expand its operation into foreign markets. This is a reason for foreign direct investment.

Despite pointing out the importance of the ownership advantage for FDI, Hymer (1976) and Kindleberger (1969) do not explain how multinational firms (MNFs) may benefit from such an advantage (Agarwal, 1980; Rugman, 1986). This point is addressed in the theory of internalisation proposed by Buckley and Casson (1976) which will be discussed below. However, firstly, this study reviews some major theoretical approaches to the debate on the theory of FDI.

Apart from the theory of market power proposed by Hymer and by Kindleberger, Vernon (1966, 1971) used the concept of the product life cycle to explain FDI. Vernon suggested that the production of a commodity goes through three distinct stages, including the ‘new’, then the ‘mature’, and finally to the ‘standardised’ commodity. In the first stage when the product is new, it is firstly designed and manufactured in home developed markets whose infrastructure and market conditions can facilitate the innovation of new products. The second stage is when the product is maturing, the designs of new products become accepted and the production process is stabilised. At that time, demand would develop for the product in overseas markets where high-income customers welcome innovation and are willing to pay a high price for it. Therefore, firms should expand their sales by exporting their commodities to other developed countries whose consumers have similar purchasing power to that of the home country.

Finally, when the product is standardised in its production, technological inputs and market knowledge are not very important. At that time, firms search for lower-cost locations abroad, particularly in less developed countries, in order to obtain cost advantages. At this stage, the product is manufactured in the less developed countries to serve their domestic consumers and to export back to the home countries and other developed countries. The firm may thus be able to increase its market share.

However, the theory of a product life cycle is mainly restricted to industries characterised by a high level of innovation (Solomon, 1978). In addition, this theory most likely addresses the position of US firms in the 1950s and 1960s when they were leaders in production innovation. Today new products are introduced at the same time in many different countries and production facilities can be located in many countries right from the beginning, because the technology and income gap between the US and other countries has narrowed since the 1970s (Moosa, 2002). Therefore, this theory is likely to be of lesser importance in the explanation of FDI activities of firms today (Giddy, 1978; Clegg, 1987).

Closely related to the product-life-cycle theory suggested by Vernon is the oligopolistic-reaction theory proposed by Knickerbocker (1973) which considers FDI as the response of a mature firm in an oligopolistic market to its competitors' decision to carry out direct investment overseas. In an oligopolistic environment, firms follow each other into foreign markets as a defensive strategy, because the firm that takes the first step in a new market exploiting any business opportunity draws the attention of similar firms that may exploit the same opportunities. However, the theory is sometimes said to be limited in explaining FDI, because it can only explain why oligopolistic firms invest defensively to counter the FDI of the initiating firm, but cannot explain the investment made by the initial firm.

A theory of currency area explains FDI based on the role of fluctuations of the exchange rate. This theory gives two different explanations of the effect of the exchange-rate fluctuations on FDI. The first argues that the exchange rate is often volatile, thus firms seek FDI to avoid the volatility of the exchange rate (Aliber, 1970; Cushman, 1985). A country with a high variation of its exchange rate may see an increase in inward FDI. In contrast, Kohlhagen (1977) and Benassy-Quere *et al.* (2001) argue that a host country with large fluctuations of the exchange rate may deter inward FDI because investors worry that these fluctuations may lead to uncertainty over the economic environment of that country.

Differentiating from the theory of currency area, Rugman (1976) and Lessard (1976) put forward another theory based on risk diversification. FDI in this theory is explained as a way for firms to spread risk from solely producing domestically. However, Caves (1988) asserts that the diversification of MNFs is more likely to result from investments that were propelled by other motives.

Unlike the theories above, Kojima (1977) argues that FDI is a means to exploit factor endowments in the host country. He states that the flow of FDI should target countries which can be assisted by the inputs of the investing firm in industries where the home country is disadvantaged. Using the case of Japan, he argues that Japanese firms tend to launch FDI in industries such as textiles, iron and steel, and assembly of motor vehicles and electronics which are less well-suited for manufacturing in Japan because of the lack of labour and resources, and strict policies on pollution. Petrochilos (1989) criticises this theory in that it is mainly relevant to the Japanese context. Thus it does not provide a general explanation of FDI. However, to some extent, Kojima's theory seems to be within the notion of locational advantages in the eclectic theory which will be discussed below.

Closely related to Kojima's theory, Helpman (1984) proposes a theory of vertical FDI. This theory explains that vertical FDI is implemented by MNFs to exploit the differences in endowments (e.g. labour costs, tax rates) between the investing country and the host country in order to decrease production costs. Hence, FDI determinants in the theory of vertical FDI are likely to be consistent with those in Kojima's theory. Differentiating from the theory of vertical FDI, Markusen (1984) puts forward a theory of horizontal FDI. According to this theory, MNFs conduct FDI to serve the local market of the host country from local production in order to save on transport costs. In this theory, the motivation for the horizontal FDI is the host country's market size and transport costs. These motives could be considered as locational factors in the eclectic theory.

Another theoretical approach to explain FDI is the theory of internalisation as suggested by Buckley and Casson (1976). Whereas Hymer (1976) and Kindleberger (1969) emphasise the importance of ownership advantages, Buckley and Casson stress internalisation advantages as an explanation of overseas investment of MNFs. The idea of internalisation theory originated from Coase (1937) who used the concept to explain the growth of multi-plant domestic firms. He argued that if transaction costs in external markets - for instance, contractual obligations or contract prices - were high, firms would internally conduct these transactions within the firm at a lower cost.

Applying Coase's internalisation approach to explain FDI, Buckley and Casson (1976) argue that firms prefer to exploit their ownership advantages such as knowledge or technology by transferring them within an internal structure (e.g. from its headquarters to subsidiaries). When the internalisation is undertaken across national borders, FDI occurs. According to Buckley and Casson (1976), the internalisation process helps investors to be able to ensure product quality as well as to keep their ownership-specific advantages within their internal firms. In addition, through the internalisation, MNFs may avoid time lags and high transaction costs.

In general, along with the theory of market power suggested by Hymer and Kindleberger, internalisation theory offers an insight into the operations of MNFs. However, it cannot explain fully the aspects of FDI as a general theory (Parry, 1985; Dunning, 1988). Theories of market power and internalisation seem to be able to explain only why a firm seeks FDI (because it possesses one or some ownership-specific advantages) and how it can exploit ownership advantages (by internalisation), but cannot fully explain why the distribution of FDI varies across countries. In other words, the theories are likely to be unable to provide an 'explicit' explanation regarding the location of FDI. This is addressed by the eclectic theory suggested by Dunning (1981, 1988) which is presented below.

The eclectic theory combines ownership, internalisation advantages and locational advantages within a single paradigm in order to interpret the main influences on FDI. According to the eclectic theory, for a firm to engage in FDI activities, the decision problem needs to satisfy the three following conditions. Firstly, a firm must possess certain advantages that provide it with comparative advantages in the host market. These advantages largely take the form of intangible assets (e.g. knowledge or technology) that are exclusive or specific to the firm possessing them, which are called ownership-specific advantages. Secondly, assuming a firm possesses one or some ownership-specific advantages, it must be more efficient for the firm to internally exploit its specific ownership advantages overseas by itself, rather than to sell them to foreign firms through market transactions. This is called an internalisation advantage, which explains how a MNF can exploit the profitability from their ownership-specific advantages. Thirdly, the host country must possess location-specific advantages that help firms to be able to make profits when operating there. The locational advantages can explain the location of FDI.

Among the theories of FDI, the eclectic theory is widely accepted as a general theory of FDI because it synthesises different theories of FDI (Dunning, 1992; Moosa, 2002). The eclectic theory encompasses ownership advantages in Hymer (1976) and Kinderberger (1969), the process of internalisation in Casson and Buckley (1976) and location-specific advantages including FDI determinants suggested in Kojima’s theory, theory of currency area, and theories of vertical and horizontal FDI. Therefore, it can give a comprehensive explanation for many aspects of FDI activities. In terms of determinants of FDI, it can be seen that the ‘original’ factors determining FDI in the perspectives of ownership advantages and internalisation advantages are likely to be similar. They often are ownership-specific factors such as advanced technology or superior managerial systems, whereas locational advantages refer to factors relevant to the host countries' characteristics, for instance, market size or labour costs. A summary of the theories reviewed and their connections is presented in Table I.2.1.

Table I.2.1: Summary of theories and their connections

Theory	**Determinants**	**Link to other theory**
Market-power theory	Ownership advantages (e.g. intangible assets such as knowledge, technology)	Internalisation theory, eclectic theory.
Product-life-cycle theory	Developmental comparative advantages	No
Oligopolistic-reaction theory	Follow the leader's investment behaviours.	No
Currency area theory	Exchange-rate fluctuations	Eclectic theory
Risk diversification theory	Risk diversification	No
Kojima's theory	Relative endowments	Vertical FDI theory, eclectic theory
Vertical FDI theory	Relative endowments	Kojima's theory, eclectic theory
Horizontal FDI theory	Market size and transport costs	Eclectic theory
Internalisation theory	Intangible assets (e.g. knowledge, technology)	Market-power theory, eclectic theory
Eclectic theory	Ownership/internalisation advantages and location-specific advantages.	Market-power theory, currency-area theory, Kojima's theory, vertical FDI theory, horizontal FDI theory, internalisation theory.

In the empirical literature on FDI determinants, macro-panel analyses often include location-specific factors, or more specifically host-country factors, rather than ownership-specific factors, to explain the variation of FDI across countries. This is so because the characteristics of the host country play key roles in the location of FDI. Hence, the current research concentrates on locational factors to explain the variation in US FDI across OECD countries. The focus on locational determinants also arises from the difficulty in measuring correctly ownership-specific factors at the country level. In the eclectic theory, though there are many locational factors that may determine FDI, which ones are important remains an empirical matter. Given a number of potential locational determinants of FDI, the research in this thesis focuses on factors which are widely included in empirical studies. The factors that are used in this study are as follows.

Market size

The size of the host country's market is generally considered as a potential locational factor determining FDI. Multinational firms often choose to invest in a country whose market is large enough, so that their turnover can exceed, at least, various costs of operating in an unfamiliar market (Davidson, 1980; Nigh 1985). A large market size of the host country can provide investors with the opportunity to capture economies of scale and to increase their profit (Scaperlanda and Mauer, 1969; Agarwal, 1980). However, evidence in some studies such as Clegg and Scott-Green (1998) and Yang *et al.* (2000) indicates an insignificant association between the host country's market size and FDI. Therefore, the hypothesis of a significant and positive association between the size of the host country's market size and FDI is not always supported in the empirical literature. This study includes market size in the model to test its effect on US FDI to OECD countries in the period 1982-2010.

Relative tax rates

Along with the size of the market, the taxation of a country commonly appears as a potential factor that may impact on FDI. A country with high tax rates may deter investors from locating their FDI there because the high tax rates can increase their costs and decrease their after-tax profits. Thus, the tax rates of the host country are expected to influence FDI negatively. However, evidence in some empirical works such as Wheeler and Mody (1992) and Swenson (1992) does not support the hypothesis of a negative relationship between tax rates and FDI. The model in this study takes relative tax rates between the host country and the US into account to test the impact of relative tax rates on FDI from the US to the host country for 1982-2010.

Relative labour costs

Labour costs frequently play an important role in determining FDI. Lower labour costs can help a firm to reduce its operation costs and production costs and thereby increase its profit. Therefore, higher labour costs in the host country relative to the investing country may lead to a decrease in FDI from the investing country to the host. Empirical evidence, however, does not always support the hypothesis of a negative relationship between labour costs and FDI in recipient countries. For example, works by Koechlin (1992) and Loree and Guisinger (1995) find an insignificant relationship between labour costs and US FDI. This study controls for relative labour costs between recipient countries and the US in the model and examines its influence on US FDI to the recipients.

Relative skilled labour

Along with labour costs, the availability of skilled labour is commonly suggested as a potential factor that may determine the location of FDI. When

MNFs establish affiliates in a foreign location they often bring knowledge and technology, which may require skilled labour in the location where they operate. Therefore, a country with skilled labour in abundance may attract more inflows of FDI, other things equal. In the empirical literature, labour skills are often measured by the gross secondary school enrolment rate or the literacy rate. Empirical studies such as Narula (1996) and Noorbakhsh *et al.* (2001) show a significant and positive relationship between the skilled labour endowment and inward FDI, while other studies, for instance, Schneider and Frey (1985) and Wei (2000), find that this relationship is insignificant. In this study, we take account of the skilled labour abundance of the recipient country relative to that of the US to investigate its impact on FDI from the US to the recipient.

Openness

Besides market size, tax rates, labour costs and skilled labour abundance, the openness of the host country is frequently mentioned as a potential factor that may affect the FDI decision-making of MNFs. Openness here is often a measure of the degree of openness of a country to international business. In empirical studies, the influence of openness on overseas direct investment seems to be ambiguous. Studies such as Culem (1988) and Moosa and Cardar (2006) find that the effect of openness on FDI is significant and positive while other studies such as Schmitz and Bieri (1972) and Wheeler and Mody (1992) find that this effect is insignificant. In this current study, we try to investigate the influence of the openness of the host country on US FDI to OECD countries. In view of the mixed results of the existing literature, the study tries to shed light on the significance of openness in the US-OECD context.

Fluctuations of the exchange rate

Another factor often considered as a potential factor which may influence FDI is the fluctuations of the exchange rate. Yet the empirical literature provides mixed results on the association between exchange rate fluctuations and FDI. Some empirical studies, for example those by Cushman (1988) and Goldberg and Kolstad (1995), find that the effect of exchange rate variability on FDI is significantly positive, while other studies, for instance those by Itagaki (1981) and Benassy-Quere *et al.* (2001), find that the fluctuation of the exchange rate has a significantly negative influence on FDI. Some other studies, such as those by Gorg and Wakelin (2002) and Crowley and Lee (2003), report an insignificant association between exchange rate variability and FDI. The inclusion of the fluctuations of the exchange rate as a driver for FDI is also controversial from a theoretical point of view. FDI is a long-term investment while exchange rate fluctuations are short-term. Furthermore, over the last thirty years or so, financial markets have become quite sophisticated in that exchange rate risk can often be hedged at relatively low cost. Thus, whether exchange rate variation is still an important variable in determining FDI is largely an empirical matter. This study controls for the fluctuation of the exchange rate by including a measure of exchange rate variability in the model to check for its effect on FDI.

Transport costs

Apart from the factors above, transport costs may be another factor affecting FDI. This is because if transport costs between the investing country and the host country are high, firms may switch from exports to FDI to serve the host country's market from local production. Hence, transports costs are included in the model to investigate their effect on FDI from US to the host countries.

Political risks

Additional factors such as political risks are likely to be a potential factor which investors consider carefully before making a FDI decision. These risks are commonly related to the investment climate, corruption, internal conflicts, ethnic or religious tensions, external conflicts and the legal environment in the host country. Since firms tend to avoid uncertainty and risks, a host country with a high extent of political risk may discourage investors. Based on different proxies, some empirical studies, for example those by Schneider and Frey (1985), Nigh (1985), Lee and Mansfield (1996) and Janicki and Wunnava (2004), provide evidence that political risks significantly influence FDI, while some others, such as those by Bennett and Green (1972), Wheeler and Mody (1992) and Bevan and Entrin (2004), find little evidence for a correlation between political risks and FDI. In this study, since we are interested in FDI in developed countries where risks relevant to major ethnic or religious tensions, and severe external or internal conflicts are comparatively rare, these particular factors are excluded from the model. Nevertheless, there may remain some political risks, and the model takes account of the investment climate, corruption and the legal environment of the recipient country in order to check the effect of these risks on FDI from the US.

Above, this study discussed the FDI determinants to be investigated in the empirical analysis. Although all these observed variables are locational factors in the eclectic theory, some of them are also determinants in other theories. In particular, the host country's market size and transport costs are considered as FDI determinants in the theory of horizontal FDI. Relative labour costs and relative skilled labour are factors relevant to Kojima's theory while relative labour costs, relative skilled labour and relative tax rates are factors relevant to the theory of vertical FDI. Fluctuations of the exchange rate are the FDI determinant suggested in the theory of currency area. A summary of the observed variables and relevant theories is presented in Table I.2.2 below.

Table I.2.2: Summary of observed variables and relevant theories

Variable	**Theory**
Market size	Eclectic theory, horizontal FDI theory
Relative tax rates	Eclectic theory, vertical FDI theory
Relative labour costs	Eclectic theory, Kojima's theory, vertical FDI theory
Relative skilled labour	Eclectic theory, Kojima's theory, vertical FDI theory
Openness	Eclectic theory
Fluctuations of exchange rate	Eclectic theory, currency area theory
Transport costs	Eclectic theory, horizontal FDI theory
Political risks	Eclectic theory

CHAPTER I.3:

METHODOLOGY

This chapter includes two main sections and aims to discuss the model, data and empirical methods used in this study. The first section presents the model and data sources while the second section discusses the main empirical methods used in this study.

I.3.1. The empirical model and data

The discussion of the determinants of FDI in the previous chapter suggests the following possible relationship:

$$\mathrm{FDI}_{\mathrm{it}} = f(\mathrm{GDP}_{\mathrm{it}}, \mathrm{TAX}_{\mathrm{it}}, \mathrm{COST}_{\mathrm{it}}, \mathrm{SKILL}_{\mathrm{it}}, \mathrm{OPEN}_{\mathrm{it}}, \mathrm{FER}_{\mathrm{it}}, \mathrm{TC}_{\mathrm{it}}, \mathrm{RISK}_{\mathrm{it}}) \quad (\mathrm{I.1})$$

where *i* and *t* denote FDI-recipient country and time indexes respectively. FDI denotes US foreign direct investment in each recipient country; GDP proxies for the recipient country's market size, measured as total output. TAX, COST and SKILL denote the relative tax rates, relative unit labour costs and relative skilled labour abundance between the recipient country and the US respectively. TC denotes transport costs between the US and the recipient country. OPEN, FER and RISK denote the recipient country's openness, fluctuations of the exchange rate and political risks respectively.

The dependent variable in this study is measured by the real US FDI stock in each recipient country. Data on the nominal US FDI stock are from the BEA, and are converted into constant 2005 US dollars (in millions) using the GDP deflator. The latter is from the International Financial Statistics (IFS) provided by the International Monetary Fund (IMF). Data on the US FDI stock from the BEA are available from 1982 onwards, only.

Real gross domestic product is used as a measure of the size of the host country's market. Data on GDP are collected from the World Economic Outlook of the IMF, and then are converted into constant 2005 US dollars (in millions) by using GDP deflators and corresponding exchange rates. Exchange rates are collected from the IFS.

We use the corporate income tax rate as a proxy for the tax rate in order to construct the relative tax rates between the host country and the US. Data on corporate income tax rates of the US and host countries were collected from the Centre for Tax Policy and Administration of the OECD.

The variable COST measures unit labour costs of the host country relative to those in the US. Data on unit labour costs were collected from the statistics database of the OECD. The relative skilled labour is proxied by the ratio of the secondary gross enrolment rate of the host country to that of the US in this study. Data on the secondary gross enrolment ratio of the US and recipients were collected from the WDI. The openness of the host country is measured by the ratio of exports to gross domestic product. Data on openness were obtained from the WDI.

We use the standard deviation of the real exchange rate as a proxy for the extent of exchange rate fluctuations. Data on the nominal exchange rate of the recipient country's currency against the US dollar were collected from the IFS, and then converted into a real exchange rate using GDP deflators. For Euro-area countries, exchange rates before 1999 were calculated based on the conversion rate between the Euro in 1999 and the country's currency. For example, for France, the exchange rate of the Euro against the US dollar in, say, 1990 is calculated by the 1990 Franc/USD exchange rate divided by the fixed conversion rate of Francs to Euros in 1999.

This study uses the CIF/FOB ratio reported in the Direction of Trade Statistics of the IMF as a proxy for transport costs. This ratio gives the value of imports including costs, insurance and freight (CIF) relative to their free on board (FOB) value, and thus it can reflect transport costs (for example, Limao and Venables, 2001; UNCTAD, 2006). With respect to political risks, this research constructs an index as the sum of ratings of the investment profile, corruption, and law and order provided by the International Country Risk Guide (ICRG). In the ICRG, the rating of investment profile of a country is scaled from 0 (very high risk) to 12 (very low risk), the rating of corruption from 0 (very high risk) to 6 (very low risk), and the rating of law and order from 0 (very high risk) to 6 (very low risk). We give equal weights to the investment profile, corruption, and law and order in the index by converting their ICRG ratings into a scale from 0 (very high risk) to 10 (very low risk) before taking the sum. This leads to an index ranging from 0 (very high risk) to 30 (very low risk). This index is used as a proxy of risks relevant to FDI decisions.

It is worth noting that some of the literature analyses additional (locational) variables that are not included in equation (I.1). These could be cultural differences, geographic distance and language differences between the host country and the investing country, all of which may affect FDI. These long-term factors are likely to be constant or approximately constant over time and will be treated as time-invariant, country-specific (fixed) effects in the empirical analysis.

This study uses a sample comprising twenty one developed OCED countries covering the period from 1982 to 2010. The countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland and the United Kingdom. In some cases in the sample where data are missing, missing data were interpolated based on available data.

Model (1) is conventionally expressed in multiplicative form as:

$$\mathrm{FDI}_{\mathrm{it}} = \mathrm{GDP}_{\mathrm{it}}^{\beta_{1i}}\mathrm{TAX}_{\mathrm{it}}^{\beta_{2i}}\mathrm{COST}_{\mathrm{it}}^{\beta_{3i}}\mathrm{SKILL}_{\mathrm{it}}^{\beta_{4i}}\mathrm{OPEN}_{\mathrm{it}}^{\beta_{5i}}\mathrm{FER}_{\mathrm{it}}^{\beta_{6i}}\mathrm{TC}_{\mathrm{it}}^{\beta_{7i}}\mathrm{RISK}_{\mathrm{it}}^{\beta_{8i}}\exp(\varepsilon_{\mathrm{it}})$$

where the βs denote elasticities and $\varepsilon_{\mathrm{it}}$ denotes the error term. $(\mathrm{I}.2)$

Taking the natural logarithms of equation (I.2) yields a log-linear form as follows:

$$\ln \mathrm{FDI}_{\mathrm{it}} = \beta_{1i} \ln \mathrm{GDP}_{\mathrm{it}} + \beta_{2i} \ln \mathrm{TAX}_{\mathrm{it}} + \beta_{3i} \ln \mathrm{COST}_{\mathrm{it}} + \beta_{4i} \ln \mathrm{SKILL}_{\mathrm{it}}$$
$$+\beta_{5i} \ln \mathrm{OPEN}_{\mathrm{it}} + \beta_{6i} \ln \mathrm{FER}_{\mathrm{it}} + \beta_{7i} \ln \mathrm{TC}_{\mathrm{it}} + \beta_{8i} \ln \mathrm{RISK}_{\mathrm{it}} + \varepsilon_{\mathrm{it}} \quad (\mathrm{I}.3)$$

where ln denotes natural logarithm. This log-linear form allows us to interpret the coefficients as elasticities. In addition, it could help to reduce the potential problem of heteroscedasticity in the error variance.

Model (I.3) is a static model. However, information on FDI determinants often becomes available with a lag relative to the time of the investment decision. In addition, there may be an additional lag from the decision-making process to actual FDI. Therefore, the effects of explanatory variables in the model (I.3) could be expected to appear with a delay. For example, in year t investors intend to invest overseas while the available information is from the previous year. Furthermore, the decision making process and the preparation for FDI such as the mobilization of funds, the building of partners, negotiations with the host country, etc., will take additional time. Overall, it may take two years or more before FDI is carried out in the host country. Hence, we experiment with two possible lag lengths, of one and two years. It is thought that a two-year lag is likely to be sufficient, considering that the data are annual and the sample size is relatively small. Even though there may be common-sense reasons for the variables to have delayed effects, there is no formal theory of dynamic adjustment in the literature and the determination of the appropriate lag length is an empirical matter. Further, lag specifications for the explanatory variables would also be useful in order to avoid the potential simultaneous influence of the dependent variable on the explanatory variables.

It may also be argued that leads (i.e., forward-looking variables) could be included to extend the dynamics in order to account for the role of expectations. For instance, it is reasonable to assume that investors may decide to invest in a country not only on the basis of past growth, but perhaps even more so based on future expected growth, since the latter may be regarded as more important for the success of the investment. However, including expectations of future growth is difficult in this setting due to the lack of availability of data on expected values of variables in the model for the entire group of countries over the period. Including leads, instead, may approximate expectations to some degree. However, the implication of the lead is perfect foresight which appears to be a rather strong assumption for the variables within the model. For instance, forecasts for GDP can be unreliable, particularly at long horizons, which are relevant for FDI decisions. Therefore, no experiment with the use of leads was attempted. This is also consistent with most of the existing literature on FDI.

I.3.2. Empirical methods

Before discussing estimation methods, this study discusses three distinct unit root tests that will be applied in the empirical analysis to check for stationarity. This is necessary because the time-series dimension, *T*, of the sample used in this research is fairly large. The problem with thin and long panel data sets is that regression results may be spurious when variables are non-stationary.

Therefore, as a first step, variables need to be tested for stationarity and, should they be non-stationary, the relationship between them needs to be tested for cointegration. Only when there is cointegration can inferences reliably be made. Otherwise, the results may be spurious.

There are various panel unit-root tests in the econometric literature, of which the LLC test proposed by Levin and Lin (1992) and Levin, Lin and Chu (2002) is a popular one (Baltagi, 2008). The null hypothesis in this test is that all panels have a homogeneous unit root versus the alternative hypothesis that all panels are stationary. In comparison with other homogeneous panel unit-root tests such as the one by Harris and Tsavalis (1999), the LLC test is likely to be more appropriate for this research because it requires the time-series dimension of the dataset to be larger than the cross-section dimension.

The potential disadvantage of the LLC test is that it restricts all autoregressive coefficients to be homogeneous across all panels. This assumption may be too strong. Maddala and Wu (1999) propose a panel unit-root test (henceforth the M-W test) that allows the autoregressive coefficients to vary across panels. In particular, this test combines the significance levels of individual Phillips-Perron or ADF unit-root tests for each cross-section i to construct an overall test statistic based on a test suggested by Fisher (1932):

$$\lambda = -2 \sum_{i=1}^{N} \ln \varphi_{\mathrm{i}} \tag{I.4}$$

where φ_{i} is the p-value of a unit root test for country *i*.

This is used to test the null hypothesis that all panels have a unit root versus the alternative hypothesis that at least one panel is stationary. Since $(-2 \ln \varphi_{\mathrm{i}})$ is distributed as χ^2 with two degrees of freedom, λ has a χ^2 distribution with 2*N* degrees of freedom where *N* denotes the number of panels.

Note that both the LLC and M-W tests are based on the potentially restrictive assumption that individual time series in the panel are cross-sectionally independent. Pesaran (2007) suggests a test that relaxes this assumption (henceforth the CIPS test) which controls for the possible presence of cross-section dependence. The null hypothesis in this test is that all panels (here, countries) have a unit root against the alternative hypothesis that a fraction of panels are stationary. In particular, the method of this test is based on augmenting the usual ADF regression with the cross-section averages of lagged levels and first-differences of the individual series to capture cross-sectional dependence. Pesaran calls this a cross-sectionally augmented Dickey-Fuller (CADF) test. The simple CADF regression is:

$$\Delta z_{it} = a_i + \rho_i z_{i,t-1} + b_0 \bar{z}_{t-1} + b_1 \Delta \bar{z}_t + \varepsilon_{it} \qquad (\mathrm{I.5})$$

where $\bar{z}_t$ is the cross-section average of z_t at time *t*. The presence of the lagged cross-section average and its first-difference can account for cross-section dependence. In the case that there is serial correlation in the errors, the regression is additionally augmented with the lagged first-differences of both z_{it} and $\bar{z}_t$ to control for serial correlation, which leads to

$$\Delta z_{it} = a_i + \rho_i z_{i,t-1} + b_0 \bar{z}_{t-1} + \sum_{j=0}^{p} b_{j+1} \Delta \bar{z}_{t-j} + \sum_{k=1}^{p} c_k \Delta z_{i,t-k} + \varepsilon_{it} \qquad (\mathrm{I.6})$$

After performing the CADF regression for each cross section, the CIPS test averages the *t*-ratio of the lagged value (henceforth $CADF_i$) to construct the *CIPS*-statistic as follows:

$$CIPS - \text{statistic} = \frac{1}{N} \sum_{i=1}^{N} CADF_i \qquad (\mathrm{I.7})$$

Pesaran (2007) also shows that the CIPS panel unit-root test has satisfactory size and power even for relatively small values of the cross-section dimension *N* and time-series dimension *T*. Along with the LLC and M-W tests, the CIPS test is used to check for unit roots in the variables used in this study. Since the data on each variable used are yearly (not daily or monthly) and the time series dimension, *T*, is not very large, the maximum lag length is chosen to be three. Among the three unit-root tests, the CIPS approach is preferred because it allows for the heterogeneity of autoregressive coefficients across panels and can address cross-sectional dependence.

Next, we turn to the discussion of the estimation methods. In this section, estimation methods are discussed that address major potential problems of this study. Firstly, the panel data set has a reasonably long time dimension and thus non-stationarity of the variables in the model needs to be addressed. In addition, since this is a macroeconomic panel data study on FDI, where many of the determinants (as discussed in the literature review) cannot be included due to data availability, these effects need to be controlled for to avoid omitted variable bias. Also, we discuss methods to allow the effects of explanatory variables and unobserved common factors to vary across countries to fulfil the aims of this study.

Consider a form of an FDI model as follows:

$$y_{it} = \beta_i^{'} x_{it} + \varepsilon_{it} \tag{I.8}$$

where y is the dependent variable, x is a vector of observed explanatory variables and β are the slope parameters for the elements of x. In this study, the explanatory variables are assumed to be exogenous.

In order to account for common factors such as a global financial crisis, this study introduces unobserved common factors (w_t) into the model (I.8). In the context of investigation of FDI from the US to a cross-section of OECD countries, w_t is also assumed to include the (average) ownership advantages of US firms such as advantages in technology, innovation or superior managerial skills which may affect US FDI to a cross section of OECD countries. The model is:

$$y_{it} = \beta_i^{'} x_{it} + \gamma_i^{'} w_t + \varepsilon_{it} \tag{I.9}$$

where w_t is assumed to be one or more latent factors capturing the effect of unobserved common factors, and γ_i are the factor loadings which may vary across countries.

The formulation in (I.9) is sometimes called an interactive fixed effects specification, and it generalizes conventional country-specific (fixed) effects and conventional time dummies (e.g. if one of the w_t factors is constant over time, that yields a set of country-specific effects; and if one of the factors has the same coefficients, that yields time effects, given that the time path of w_t is not restricted). Therefore, model (I.9) can control for country-specific effects (e.g. cultural differences, geographic distances) and conventional time dummies, but is more general than either.

Now we discuss the estimation of model (I.9) by using different estimators. Firstly, it can be seen that the Pooled OLS (POLS) estimator uses a conventional least squares regression based on pooling all the observations without considering country-specific effects, which could lead to biased estimates. In addition, in the POLS estimates, the effects of the explanatory variables (x) are restricted to be constant across countries ($\beta_i = \beta$). Unobserved common factors (w_t) might be taken into account by introducing time dummies into the POLS regression model. However, the time dummies can only capture common shocks to FDI that have the same effects across countries, and thus the effects of w_t on FDI are constrained to be homogeneous across countries $(\gamma_i = \gamma)$ in the POLS estimates.

In the Fixed Effects (FE) and Random Effects (RE) estimators, time-invariant country-specific effects are taken into account and treated as fixed and random in the regression respectively. To decide between the FE and the RE estimator, we can run a Hausman test where the null hypothesis is that the preferred model is RE versus the alternative being FE. However, in the FE and RE estimators, the slope parameters of x are constrained to be identical across countries $(\beta_i = \beta)$. In addition, as in Pooled OLS estimation, w_t may be taken into account by including time dummies in the FE and RE regression models, and thus the influence of w_t on FDI is restricted to be constant across countries $(\gamma_i = \gamma)$ by both the FE and RE estimators.

Apart from the POLS, FE and RE estimators, the more recent Mean Group (MG) estimator, proposed by Pesaran and Smith (1995), relaxes the assumption of homogeneity of explanatory variables' parameters. The MG estimator allows the effects of explanatory variables to vary across countries by firstly estimating country-specific OLS regression models and then averaging the estimated parameters across countries to obtain an average effect. In addition, this estimation can capture country-specific effects as an intercept in each of the individual regression models (one per country). Unobserved common shocks (w_t) may be controlled for by introducing a time trend in the regression model for each country, and thus the effect of w_t is allowed to vary across countries (Eberhardt and Bond, 2009; Eberhardt, 2011). However, the use of the country-specific time trend will restrict the unobserved common factors to be (smoothly) increasing or decreasing over time. Note that we cannot use a full set of time dummies (as in the POLS and FE estimators) in the regression model for each country because they would explain the dependent variable perfectly.

Recent work by Pesaran (2006), extended to non-stationary variables by Kapetanios, Pesaran and Yamagata (2011), suggests the use of Common Correlated Effects (CCE) estimators with cross-section averages of the dependent variable $(\bar{y}_t)$ and independent variables $(\bar{x}_t)$ to account for the presence of unobserved common factors (w_t) with heterogeneous effects (Pesaran, 2006; Coakley, Fuertes and Smith, 2006; Kapetanios, Pesaran and Yamagata, 2011; Pesaran and Tosetti, 2011); then the model (I.9) becomes

$$y_{it} = \beta_i^{'} x_{it} + c_i \bar{y}_t + d_i^{'} \bar{x}_t + \varepsilon_{it} \qquad (\mathrm{I}.10)$$

In CCE estimates, the estimated country-specific parameters on $\bar{y}_t$ and $\bar{x}_t$ are not interpretable in a conventional way: their presence is only to control for the biasing effects of the unobserved common factors. There are two alternative methods to estimate model (I.10), namely the Common Correlated Effects Pooled (CCEP) and Common Correlated Effects Mean Group (CCEMG) estimators. Pesaran (2006), Stock and Watson (2008), Kapetanios, Pesaran and Yamagata (2011) and Pesaran and Tosetti (2011) show that the CCE estimators are robust to heteroskedasticity and serial correlation and to the presence of structural breaks. The CCEP estimator is a fixed effects regression where each country has a separate parameter for each of the cross-section averages. Therefore, the CCEP allows unobserved common factors to have heterogeneous effects across countries. However, in CCEP estimation, the parameters of the main explanatory variables (here, the *x*’s) are restricted to be identical across countries $(\beta_i = \beta)$. Alternatively, we can relax the restriction of the homogeneity of the slope parameters by using a CCEMG estimator. The

CCEMG estimator, which is based on an MG estimation of model (I.10), can permit the observed explanatory variables' parameters to be varying across countries. As with CCEP, it also allows the unobserved common factors to have different effects on different countries.

It is worth noting that if variables are non-stationary, regression results could be spurious. However, this is not the case when the variables are cointegrated. Normally, when variables are non-stationary, their linear combination is also non-stationary which undermines inference and leads to spurious regression results. However, non-stationary variables may move together over time even though individually they are random walks. In other words, cointegration is a specific result which may occur in the presence of variables with unit roots. As a result of cointegration, the error term is stationary. An empirical indicator of cointegration is when a regression produces stationary residuals. As discussed, the current study accounts for unobserved common factors in the estimation, and thus they could be a part of a cointegrating vector. Since the way to control for unobservable common factors varies across estimators, this study will first estimate the model with the inclusion of unobserved common factors, and then check for the stationarity of the residuals. If observed explanatory variables (and unobserved common factors) are cointegrated, we can establish a long-run economic relationship between the variables which can be interpreted in relation to the economic theories of FDI presented in the literature review above.

Another problem is that the observed explanatory variables and unobserved common factors may have effects on US FDI to recipient countries to different degrees. The restrictions that those effects are homogeneous across countries may cause cross-section dependence among regression errors, leading to biased estimates, especially in a panel data analysis with long *T*. Therefore, this study will check the cross-section independence of the residuals by using a cross-section dependence (CD) test suggested by Pesaran (2004). In this study, we use the unit-root and CD tests to choose the preferred empirical model.

In summary, this chapter provided a discussion of the model, data sources and empirical methods used in this study to investigate the influence of the host country factors on FDI from the United States (US) to developed OECD countries in the period 1982-2010. The empirical results will be presented in the next chapter.

CHAPTER I.4:

EMPIRICAL RESULTS

This chapter presents the empirical results of the research. In order to check the stationarity of the variables, this study plots variables and their first differences over time (see appendix I.2 and appendix I.3). This is because the appropriate critical values of the unit root test statistics depend on the deterministic terms that are included. If the unit root test does not specify the deterministic terms correctly, then this may lead to an over- or under-rejection of the null hypothesis. A straightforward way to decide on what deterministic terms should be included in the unit root test is to look at the graphs of the individual series.

The graphs in appendix I.2 show that the variables of foreign direct investment, $\ln \mathrm{FDI}$, market size, $\ln \mathrm{GDP}$, relative tax rates, $\ln \mathrm{TAX}$, relative skilled labour, $\ln \mathrm{SKILL}$, openness, $\ln \mathrm{OPEN}$, transport costs, $\ln \mathrm{TC}$, and political risks, $\ln \mathrm{RISK}$, are likely to be trended while the variables relative labour costs, $\ln \mathrm{COST}$, and fluctuations of the exchange rate, $\ln \mathrm{FER}$, are not likely to be trended. Therefore, this study adopts LLC, M-W and CIPS unit-root tests with a trend for the former and adopts those with a constant only for the latter. The p-values of the unit-root tests of all variables are reported in Table I.4.1. We can see from Table I.4.1 that the results of the LLC test reject the null hypothesis that variables $\ln \mathrm{FDI}$, $\ln \mathrm{TC}$ and $\ln \mathrm{RISK}$ have a unit root while those of the M-W and CIPS tests do not reject the null hypothesis.

With respect to the variables $\ln \mathrm{TAX}$, $\ln \mathrm{COST}$ and $\ln \mathrm{OPEN}$, the results of the LLC and M-W tests reject that these variables are non-stationary at conventional levels of significance. However the use of the CIPS test does not reject the null hypothesis that they are non-stationary. Table I.4.1 also shows that all three tests reject a unit root for the variable $\ln \mathrm{FER}$ but they do not reject for variables $\ln \mathrm{GDP}$ and $\ln \mathrm{SKILL}$. Among the three unit-root tests, the results of the CIPS are preferred because this test allows for the heterogeneity of autoregressive coefficients across panels and can control for cross-sectional dependence. Therefore, it can be seen that, apart from the variable $\ln \mathrm{FER}$, the other variables in the model are likely to be non-stationary.

Table I.4.1: Unit root tests for variables

	LLC (p-value)	M-W (p-value)	CIPS (p-value)
$\ln \mathrm{FDI}$	0.04	0.65	0.64
$\ln \mathrm{GDP}$	0.42	0.80	0.72
$\ln \mathrm{TAX}$	0.01	0.01	0.47
$\ln \mathrm{COST}$	0.01	0.01	0.43
$\ln \mathrm{SKILL}$	0.43	0.99	0.84
$\ln \mathrm{OPEN}$	0.01	0.01	0.85
$\ln \mathrm{FER}$	0.01	0.01	0.03
$\ln \mathrm{TC}$	0.01	0.29	0.85
$\ln \mathrm{RISK}$	0.01	0.62	0.31

Note: The lag length of the unit root tests is three. This study experimented with different lag lengths up to order three: the results did not change significantly.

Variables that are integrated of order one can be made stationary by taking first differences. Since the more reliable CIPS test suggests that all the variables except for fluctuations of the exchange rate, ln FER, may be non-stationary, the second step of the testing procedure is to find out whether the first differences are stationary. If this is the case, then the variables are integrated of order one, conventionally denoted as I(1). Since unit root tests above indicate that ln FER is stationary, there is no need to test the first difference of ln FER for stationarity.

The graphs in appendix I.3 show that all the first differences of variables seem to be un-trended, and thus this study runs the tests with no trend for the first differences of variables. We can see that all the results of LLC, M-W and CIPS tests in Table I.4.2 reject that the first-differences of the variables have a unit root at the one or five percent levels of significance, indicating that the first-differences of the variables are stationary. Therefore, from the results in Tables I.4.1 and I.4.2, it is likely that the variable for exchange-rate fluctuations, $\ln \mathrm{FER}$, seems to be stationary while the others in the model are potentially I(1).

Table I.4.2: Unit root tests for the first difference of variables

	LLC (p-value)	M-W (p-value)	CIPS (p-value)
$\Delta \ln \mathrm{FDI}$	0.01	0.01	0.01
$\Delta \ln \mathrm{GDP}$	0.01	0.01	0.02
$\Delta \ln \mathrm{TAX}$	0.01	0.01	0.01
$\Delta \ln \mathrm{COST}$	0.01	0.01	0.01
$\Delta \ln \mathrm{SKILL}$	0.01	0.02	0.01
$\Delta \ln \mathrm{OPEN}$	0.01	0.01	0.01
$\Delta \ln \mathrm{TC}$	0.01	0.01	0.01
$\Delta \ln \mathrm{RISK}$	0.01	0.01	0.01

Note: The lag length of the unit root tests is three. This study experimented with different lag lengths up to order three: the results did not change significantly.

Next, estimates of the model for FDI are reported. All the models assume that the explanatory variables are exogenous. This may be too strong an assumption. For instance, GDP is likely to be endogenous in a model that explains FDI. The consequence of including endogenous variables in the model will be biased and inconsistent coefficient estimates. However, as discussed in chapter I.3, the use of lagged values of the explanatory variables as instruments (for the current values) in the model could help to reduce this potential problem. This study, in turn, experimented with one- and two-year lags as instruments.[3] The POLS, FE, MG, CCEP and CCEMG results from the model using one-year lagged values for the explanatory variables are reported in Table I.4.3 while those from the model using two-year lagged values are reported in Table I.4.4.

The choices of the lag length and the estimation method have a strong impact on the estimation results. Turning to the results with the one-year lagged values depicted in Table I.4.3 first, none of the variables are significant across all estimates. There is some indication that market size, relative tax rates, relative labour costs, openness and the volatility of the exchange rate may be determinants of FDI. The results vary depending on the estimation method used. Turning to Table I.4.4, there is evidence that relative tax rates, relative labour costs, relative skilled labour, openness and political risks are associated with FDI. Again, the significance of the estimated coefficients on these variables seems to vary. However, the influence of market size and the fluctuations of the exchange rate on FDI are found to be significant when using most of the estimators.

---

[3] See pp.21-22 in chapter I.3 for a discussion of the choice of the lag.

Table I.4.3:

The estimation of the models using one-year lagged values for explanatory variables

	(1) POLS	(2) FE	(3) MG	(4) CCEP	(5) CCEMG
$\ln \mathrm{GDP}_{t-1}$	1.45*** (0.17)	2.01*** (0.69)	0.93*** (0.35)	0.36 (0.43)	1.05 (0.91)
$\ln \mathrm{TAX}_{t-1}$	-0.72*** (0.17)	-0.11 (0.23)	-0.28** (0.12)	-0.17 (0.12)	-0.08 (0.19)
$\ln \mathrm{COST}_{t-1}$	-1.98** (0.72)	0.73* (0.41)	-0.01 (0.16)	0.40* (0.24)	0.41 (0.49)
$\ln \mathrm{SKILL}_{t-1}$	1.08 (1.16)	0.75 (0.59)	0.11 (0.34)	0.01 (0.31)	-0.29 (0.51)
$\ln \mathrm{OPEN}_{t-1}$	1.66*** (0.52)	1.20* (0.58)	-0.06 (0.25)	0.11 (0.26)	-0.36 (0.28)
$\ln \mathrm{FER}_{t-1}$	-0.06 (0.09)	-0.05** (0.02)	-0.04** (0.01)	-0.02* (0.01)	-0.02 (0.03)
$\ln \mathrm{TC}_{t-1}$	-0.58 (0.44)	0.32 (0.27)	0.06 (0.13)	-0.72 (0.85)	0.10 (0.13)
$\ln \mathrm{RISK}_{t-1}$	1.84 (1.14)	0.04 (0.43)	0.31 (0.33)	-0.39 (0.26)	0.40 (0.53)
Observations	588	588	588	588	588
RMSE	0.8699	0.3542	0.1325	0.1692	0.0890

Note: *** denotes significance at the one percent level. ** denotes significance at the five percent level. * denotes significance at the ten percent level. Standard errors are reported in parentheses. In POLS and FE regressions, the reported standard errors are heteroskedasticity-robust and clustered by country. Hausman test suggests choose the FE estimator over the RE estimator (p-value=0.01). GDP denotes the host country's market size, TAX relative tax rates, COST relative labour costs, SKILL relative skilled labour, OPEN the host country's openness, FER fluctuations of the exchange rate, TC transport costs, RISK the host country's political risks. RMSE is root mean squared error. STATA output of the regressions is reported in appendix I.4.

Table I.4.4:

The estimation of the models using two-year lagged values for explanatory variables

	(1) POLS	(2) FE	(3) MG	(4) CCEP	(5) CCEMG
$\ln \mathrm{GDP}_{t-2}$	1.45*** (0.17)	1.97** (0.75)	1.02* (0.61)	0.77** (0.38)	1.53** (0.67)
$\ln \mathrm{TAX}_{t-2}$	-0.72*** (0.17)	-0.17 (0.24)	-0.08 (0.15)	-0.34*** (0.11)	-0.31 (0.19)
$\ln \mathrm{COST}_{t-2}$	-1.98** (0.73)	0.66 (0.44)	-0.40** (0.18)	0.11 (0.21)	0.30 (0.54)
$\ln \mathrm{SKILL}_{t-2}$	0.99 (1.17)	0.72 (0.68)	0.32 (0.46)	0.58** (0.27)	0.38 (0.47)
$\ln \mathrm{OPEN}_{t-2}$	1.67*** (0.54)	1.27** (0.61)	-0.31 (0.33)	0.40 (0.25)	0.12 (0.57)
$\ln \mathrm{FER}_{t-2}$	-0.07 (0.09)	-0.05** (0.02)	-0.04*** (0.02)	-0.03*** (0.01)	-0.03* (0.02)
$\ln \mathrm{TC}_{t-2}$	-0.54 (0.46)	0.42 (0.29)	0.07 (0.15)	0.03 (0.08)	0.01 (0.10)
$\ln \mathrm{RISK}_{t-2}$	1.99 (1.15)	0.30 (0.41)	0.45* (0.25)	0.27 (0.23)	0.73** (0.35)
Observations RMSE	567 0.8733	567 0.3490	567 0.1236	567 0.1332	567 0.0627

Note: *** denotes significance at the one percent level. ** denotes significance at the five percent level. * denotes significance at the ten percent level. In POLS and FE regressions, the reported standard errors are heteroskedasticity-robust and clustered by country. Hausman test suggests choose the FE estimator over the RE estimator (p-value=0.01). GDP denotes the host country's market size, TAX relative tax rates, COST relative labour costs, SKILL relative skilled labour, OPEN the host country's openness, FER fluctuations of the exchange rate, TC transport costs, RISK the host country's political risks. RMSE is root mean squared error. STATA output of the regressions is reported in appendix I.5.

In order to discriminate between the one and two-year lagged models, this study compares the root mean square error (RMSE) of the two models. Except for POLS, the RMSE of all other regressions for the model with two-year lags is smaller than that of the model with one-year lags. Thus, the fit of the model is better in the two-year lagged form. The result suggests FDI may be best explained by two-year lagged information rather than one-year lagged information. The section below concentrates on discussing estimation results for the models using two-year lagged values of the explanatory variables in Table I.4.4.

In POLS estimation with assumptions on the homogeneity of slope parameters and factor loadings for unobserved common factors, the coefficients on the variables of market size, $\ln \mathrm{GDP}_{t-2}$, the relative tax rates, $\ln \mathrm{TAX}_{t-2}$, the relative labour costs, $\ln \mathrm{COST}_{t-2}$ and the host country's openness, $\ln \mathrm{OPEN}_{t-2}$, are significant; the elasticities are 1.45, -0.72, -1.98 and 1.67, respectively. This result implies a one-percent increase in the market size and openness of the host country, on average, increases the level of the US FDI stock in the host country by 1.45 and 1.67 per cent, respectively, while a one-percent increase in relative tax rate and relative labour costs, on average, decreases the level of the US FDI stock in the host country by 0.72 and 1.98 percent, respectively. Other variables - the relative skilled labour, $\ln \mathrm{SKILL}_{t-2}$, the fluctuations of the exchange rate $\ln \mathrm{FER}_{t-2}$, transport costs, $\ln \mathrm{TC}_{t-2}$, and the host country's political risks, $\ln \mathrm{RISK}_{t-2}$, are found to be insignificant in the OLS estimation. Note that the POLS estimator does not control for unobserved country-specific effects, for example, cultural differences or geographic distance between the investing country and host countries, which may influence US FDI to the host country. In addition, the results of the CIPS and LLC tests (see Table I.4.5, p.40) show that the residuals estimated by POLS may contain a unit root. The implication is that the variables are not cointegrated and that the regression may be spurious. A further point is that the result of Pesaran (2004)'s cross-section dependence (CD) test (see Table I.4.6, p.41) indicates that the POLS residuals are cross-sectionally dependent. Therefore, the POLS estimation results are likely to be biased.

In order to control for country-specific effects, we can use FE and RE estimators where country-specific effects are taken into account and treated as fixed and random parameters in the regression respectively. To decide between the FE and the RE estimator, this study runs a Hausman test where the null hypothesis is that the preferred model is RE versus the alternative being FE. The result of the Hausman test rejects the null hypothesis at the one percent level of significance, implying that the FE model should be preferred over the RE model.

The results of FE estimation in Table I.4.4 show that the coefficients on variables for the host country's market size, $\ln \mathrm{GDP}_{t-2}$, the host country's openness, $\ln \mathrm{OPEN}_{t-2}$, and the variability of the exchange rate, $\ln \mathrm{FER}_{t-2}$, are significant and the elasticities are approximately 1.97, 1.27 and -0.05 respectively. These results imply that a one percent increase in the host country's market size and openness, on average, increases the level of the US FDI stock in the host country by 1.97 and 1.27 percent respectively while a one percent increase in fluctuations of the exchange rate, on average, decreases the level of the US FDI stock in the host country by 0.05 per cent. The coefficients on the other variables including relative tax rates, $\ln \mathrm{TAX}_{t-2}$, relative labour costs, $\ln \mathrm{COST}_{t-2}$, relative skilled labour, $\ln \mathrm{SKILL}_{t-2}$, transport costs, $\ln \mathrm{TC}_{t-2}$, and political risks, $\ln \mathrm{RISK}_{t-2}$, are found to be insignificant. However, similar to the POLS estimation, the results of the M-W and CIPS tests (see Table I.4.5, p.40) do not reject the hypothesis of the presence of a unit root in the FE residuals, implying that the FE regression may be spurious. In addition, the residuals estimated from the FE estimator are found to be cross-sectionally dependent on the basis of Pesaran's CD test (see Table I.4.6, p.41). Thus, the FE coefficients are likely to be biased.

In the MG estimation which allows the effects of the observed explanatory variables to vary across countries, the coefficients on the variables for market size, $\ln \mathrm{GDP}_{t-2}$, and political risks, $\ln \mathrm{RISK}_{t-2}$, are significant at the ten percent level with values 1.02 and 0.45 respectively whereas those on the variables for relative labour costs, $\ln \mathrm{COST}_{t-2}$ and the fluctuations of the exchange rate, $\ln \mathrm{FER}_{t-2}$ are approximately -0.40 and -0.04, and are significant at the one and five percent level respectively. In contrast to the results in the POLS and FE estimates, the coefficient on the variable for the host country's political risks, $\ln \mathrm{RISK}_{t-2}$, in the MG estimates is found to be significant at the ten percent level and approximately 0.45. These results imply that a one percent increase in the market size and the risk index of the host country, on average, increases the US FDI stock by 1.02 and 0.45 percent respectively, while a one percent increase in relative labour costs and the fluctuations of the exchange rate, on average, decreases the FDI stock by 0.40 and 0.04 percent respectively. Other variables, including relative tax rates, $\ln \mathrm{TAX}_{t-2}$, relative skilled labour, $\ln \mathrm{SKILL}_{t-2}$, the host country's openness, $\ln \mathrm{OPEN}_{t-2}$ and transport costs, $\ln \mathrm{TC}_{t-2}$, are found to have an insignificant effect on US FDI stock to the host country. Unlike the POLS and FE cases, all three unit-root tests suggest that the MG residuals are likely to be stationary (see Table I.4.5, p.40). This implies that there exists a cointegrating long-run relationship between the variables in the model. The CD test does not reject the null hypothesis of the absence of cross-section dependence in the MG residuals (see Table I.4.6, p.41). However, the p-value in the CD test is just 0.12, and thus the absence of cross-section dependence in the MG residuals seems not to be safely confirmed. Note that, in the MG estimation, unobserved common factors are controlled for by introducing a time trend in the regression model for each country, and thus the effect of unobserved common factors is allowed to vary across country. However, the use of the country-specific time trend restricts the unobserved common factors to be (smoothly) increasing or decreasing over time.

Table I.4.5: Unit root tests for the estimated residuals

	POLS	FE	MG	CCEP	CCEMG
LLC test (p-value)	0.30	0.01	0.01	0.01	0.01
M-W test (p-value)	0.22	0.11	0.01	0.01	0.01
CIPS test (p-value)	0.95	0.73	0.01	0.01	0.01

Note: The lag length of the unit root tests is three. This study experimented with different lag lengths up to order three: the results did not change significantly.

Next, this study uses the CCEP estimator in which the effects of unobserved common factors are permitted to be heterogeneous although the parameters of the explanatory variables are constrained to be identical across countries (as in POLS, RE and FE, but not MG). In the CCEP estimates, the variables for relative labour costs, $\ln \mathrm{COST}_{t-2}$, the host country's openness, $\ln \mathrm{OPEN}_{t-2}$, transport costs, $\ln \mathrm{TC}_{t-2}$, and the political risks of the host country, $\ln \mathrm{RISK}_{t-2}$, are found to be insignificant while the variables for the host-country market size, $\ln \mathrm{GDP}_{t-2}$, relative tax rates $\ln \mathrm{TAX}_{t-2}$, relative skilled labour, $\ln \mathrm{SKILL}_{t-2}$, and the exchange-rate variability, $\ln \mathrm{FER}_{t-2}$, are significant with their coefficients being approximately 0.77, -0.34, 0.58 and -0.03 respectively. The results indicate that a one percent increase in the relative tax rate and the fluctuation of the exchange rate, on average, reduces the US FDI stock by 0.34

and 0.03 per cent respectively, whereas a one-percent increase in the host country's market size and relative skilled labour, on average, raises the US FDI stock by 0.77 and 0.58 per cent respectively. Like the MG estimation, the results of LLC, M-W and CIPS unit-root tests indicate that the residuals estimated from the CCEP estimation are potentially stationary. However, the result of the CD test rejects the null hypothesis, implying that the CCEP residuals are potentially cross-sectionally dependent. Therefore, it is likely that the CCEP results could also be biased.

Table I.4.6: Cross dependence tests for the estimated residuals

	POLS	FE	MG	CCEP	CCEMG
CD test (p-value)	0.01	0.01	0.12	0.01	0.59

Note: CD test is Pesaran (2004) test with the null hypothesis of cross-section independence.

This study continues to attempt to improve on the estimation approach by using the recently-developed CCEMG estimator, which allows the effects of the observed explanatory variables and the factor loadings on unobserved common factors to vary across individual countries. The CCEMG regression shows that the coefficients on the variables of the host country's market size, $\ln \mathrm{GDP}_{t-2}$, and the host country's political risks, $\ln \mathrm{RISK}_{t-2}$, are found to be significant at the five percent level, with values approximately 1.53 and 0.73 respectively, whereas that on exchange-rate fluctuations is found to be significant at the ten percent level with a value of -0.03. These estimation results imply that a one percent increase in the market size and the risk index of the host country will, on average, increase the level of the US FDI stock in the host country by 1.53 and 0.73 per cent respectively, while a one percent increase in the fluctuations of the exchange rate will, on average, decrease the level of the US FDI stock in the host country by 0.03 per cent. Other explanatory variables including relative tax rates, relative labour costs, relative skilled labour, the host country's openness and transports costs are found to have insignificant effects on the level of the US FDI stock in the host country in the CCEMG estimation. In addition, this study runs an F-test to test the joint significance of cross-section averages of variables which are used to capture the heterogeneous effects of unobserved common factors in the CCEMG estimator. The result of the F-test shows that the cross-section averages are jointly significant at the one percent level.

The results of the LLC, M-W and CIPS unit-root tests (see Table I.4.5) indicate that the estimated residuals from the CCEMG estimation are potentially stationary. This means that variables (including unobserved common factors) are likely to be cointegrated, implying the existence of a long-run relationship in the data. In addition, the result of the CD test does not reject the null hypothesis with the p-value being 0.59 (see Table I.4.6), implying that the hypothesis that the CCEMG residuals are cross-sectionally independent is not rejected at conventional levels. These results indicate that the CCEMG estimation is to be preferred to the previous ones, because the POLS, FE and CCEP residuals may be non-stationary and/or cross-sectionally dependent. Although CD tests do not reject the presence of cross-sectional dependence in the residuals estimated from the CCEMG and MG estimators, the p-value of the CD test for the CCEMG residuals (equal to 0.59) is much larger than that of the MG (equal to 0.12). Moreover, the RMSE of the CCEMG estimator is found to be smaller than that of the MG estimator, implying that the fit of the model estimated by CCEMG is better than that of the model fitted by the MG estimator. Therefore, the CCEMG estimator is preferred. In terms of theory, it can be seen that the significance of the host country's market size, political risks, and fluctuations of the exchange rate in the CCEMG estimation supports the perspective of locational factors in the eclectic theory. Also, the significance of the market size variable for FDI supports the theory of horizontal FDI. In addition, a negative effect of exchange-rate fluctuations on FDI supports the theory of currency area, in that a host country with large fluctuations of the exchange rate deters inward FDI because investors may worry that those large fluctuations can lead to instability in the economic environment in that country.

Above, this chapter has reported empirical results where we in turn used POLS, FE, MG, CCEP and CCEMG estimators with different properties. The next chapter will provide the conclusion of this study.

CHAPTER I.5:

CONCLUSION

This study has used aggregate macro-panel data to investigate empirically the effects of market size, relative tax rates, relative labour costs, relative skilled labour, openness, fluctuations of the exchange rate, transport costs and political risks on US foreign direct investment to OECD countries in the period 1982-2010. In the study, we experimented with two groups of models, of which the first group consists of models using one-year lagged values of the explanatory variables and the second consists of models using two-year lagged values of the explanatory variables. This is because the explanatory variables are argued to react to FDI with a lag of one or two years. In addition, the use of lagged values could help to reduce the problem of a simultaneous effect of FDI on these variables. Empirical results in this study suggest using the models with two-year lags for explanatory variables, and thus the conclusion below is based on the estimation results from these models.

The empirical findings seem to reject the inferences from the POLS, FE and CCEP estimators because the estimated residuals achieved from these estimations are found to be cross-sectionally dependent and/or possibly non-stationary. It is possible that the cross-section dependence and/or non-stationarity of the residuals are potentially caused by the restrictions of the homogeneity of the slope coefficients on the observed explanatory variables, and the homogeneity of the factor loadings on the unobserved common factors. Unlike the POLS, FE and CCEP residuals, the MG and CCEMG residuals are found to be stationary and do not show serious evidence of cross-section dependence. In the MG estimation, which allows for the heterogeneity of the slope parameters on the observed explanatory variables, the market size of the host country was found to have a significant effect on US FDI stock to the host country in the period 1982-2010, and relative labour costs, fluctuations of the exchange rate and the host country's political risks were also found to have significant effects on the US FDI stock. Although the CD test does not reject the null hypothesis of the absence of cross-section dependence in the MG residuals at conventional levels, the p-value in the CD test is just 0.12; and thus the absence of cross-section dependence in the MG residuals seems not to be safely confirmed. Note that in the MG estimation, unobserved common factors are controlled for by introducing country-specific time trends, and thus the factor loadings on unobserved common factors are allowed to vary across countries. However, the use of the country-specific time trend restricts the unobserved common factors to be (smoothly) increasing or decreasing over time.

In the CCEMG estimation, which allows for the heterogeneity of the slope parameters on the observed explanatory variables and in the factor loadings on the unobserved common factors, in order to fulfil the research aims set out in the introduction, the host country's market size was found to have a significant effect on US FDI stock to the host country in the period 1982-2010, and a political risk index of the host country and fluctuations of the exchange rate were also found to have a significant effect on the US FDI stock. These results are likely to be similar to those of the MG estimation. However, in the CCEMG estimation, relative labour costs were found to have an insignificant impact on the FDI stock. The result differs from MG estimation where the relative labour costs were found to be significant. The influences of other variables, including relative tax rates, relative skilled labour, the host country's openness and transport costs on the US FDI stock were also found to be insignificant in the CCEMG estimation. In addition, the result of the F-test rejects the exclusion of the cross-section averages of variables which are used to capture the unobserved common factors with heterogeneous impacts on FDI from the CCEMG regression. Unlike the POLS, FE and CCEP cases, the inference from the CCEMG estimates should not be rejected because its estimated residuals were found to be stationary and cross-sectionally independent. Moreover, the inference from the CCEMG estimator is likely to be preferred over that of the MG because the p-value of the CD test for CCEMG residuals is larger than that for the MG residuals and additionally the fit of the CCEMG model was found to be better than that of MG.

In brief, the empirical results from the preferred CCEMG estimates indicate that US FDI seems to be attracted to host countries with a large market size, little risk in the investment climate, corruption or the legal environment, and stability of the exchange rate. The CCEMG regression can allow for common shocks, for example the global financial crisis, and/or average ownership-specific advantages of US firms such as advanced technology or superior managerial systems that could affect FDI from the US to OECD countries. This could be reasonable because US firms are known as leading firms in innovation, knowledge and ways to efficiently operate. The significance of the variable for political risks may also reinforce this, because the FDI motivated by these factors may tend to attach importance to the host countries which have high transparency and efficiency in the investment and business environment, and are known for the impartiality of the legal system and the observance of the law (e.g. commercial dispute regulations, assets and property or intellectual property laws).

In terms of theory, the evidence that the host country’s market size, fluctuations of the exchange rates and the host country’s risks of investment climate, corruption and legal environment have significant effects on FDI in this study supports the perspective of locational factors in the eclectic theory.

On the other hand, although transport costs are found to be insignificant, the finding of the significance of the host country's market size for FDI is likely to support the theory of horizontal FDI to some extent. In addition, the finding of the negative effect of exchange-rate fluctuations on FDI also supports the theory of currency area, in that a host country with large fluctuations of the exchange rate discourages inward FDI because investors may worry that those large fluctuations can lead to uncertainty or instability in the economic environment in that country. Vertical FDI theory and Kojima's theory seem not to be supported, because of the finding of insignificance for the relative-endowment variables. It is worth noting that the significance of the host country's market size is robustly positive throughout all estimators (including the POLS, FE, MG, CCEP and CCEMG) while that of fluctuations of the exchange rate is robustly negative in all estimators, excepting the POLS.

It can be seen that this study makes the following contributions to the panel literature on determinants driving FDI from a country to a cross-section of host countries. Firstly, the empirical literature tends to apply standard panel data estimators that constrain the observed explanatory variables' parameters to be homogeneous across recipients. This assumption can be too strong, because the influence of a factor on FDI may be heterogeneous for different countries. In this research, the data seem to reject empirical specifications which impose homogeneity of the slope parameters.

Secondly, the empirical results in this study show that the exclusion of unobserved common factors, or a constraint of the homogeneity of the factor loadings for the unobserved common factors, may potentially produce serious biases in the findings. It is noteworthy that, in the context of the investigation on FDI from an investing country to a cross-section of host countries, besides global shocks, common factors could include the time-varying average advanced knowledge, technology or innovation of the investing country's firms. Those factors seem to be likely to be relevant to ownership/internalization advantages that may influence FDI, as suggested in Hymer (1976), Kindleberger (1969), Buckley and Casson (1976) and Dunning (1977, 1981). Therefore, they should be carefully addressed in the estimation. This study adopted a new approach to address these two issues of the previous literature by employing the recent CCEMG estimator, in order to take unobserved common factors into account and permit heterogeneous effects of both observed variables and the unobserved common factors across recipients in the empirical estimation. The empirical results seem to support this approach rather than more standard ones and thereby indicate that this approach should be considered for future empirical analyses of the determinants of FDI. In addition, the panel dataset employed in this research is an additional contribution to the literature. To the best of our knowledge, the data used are more recent than in previous work on FDI from US to OECD countries. Thus, our findings are making use of additional information.

This study has the following limitations. Firstly, it does not deal with dynamics in detail. This is because the theoretical CCEMG set-up in Pesaran (2006) and Kapetanios, Pesaran and Yamagata (2011) does not allow for dynamics in the model. The CCEMG estimator was chosen because it allows for the heterogeneity of observed variables and unobserved common factors across countries. Standard dynamic panel estimators such as Anderson-Hsiao, Arellano-Bond[4] or dynamic FE do not allow for the heterogeneity of the effects of the observed variables and unobserved common factors across countries.

---

[4] Also, the Anderson-Hsiao and Arellano-Bond estimators are intended for short-T panels and may be inappropriate for the sample with $T > N$ as in the case of this study (Baltagi, 2008).

A second limitation is that the assumption of the exogeneity of regressors in CCEMG could be too strong. The use of two-year lagged values for independent variables in this study may help to reduce the effect of the dependent variables on independent variables to some extent; however, this may not avoid the problem completely.

This study has examined how the results vary across different estimators and models. In an area like this one, models are approximations at best. In that case, it makes sense to report several different models, and readers can then gain a sense of the different findings implied by different models, and identify those findings that are sensitive to the precise model estimated. There is a recent working paper by Chudik and Pesaran (2013) which extends the CCEMG estimator to weakly exogenous regressors and allows for dynamics in the model. However, the theoretical results of the CCEMG in that paper are currently only for the case when the dependent variable and regressors are stationary. In our case, since the dependent variable and most regressors appear to be I(1), the existing results are not applicable, and otherwise the theoretical properties of the CCEMG estimator applied to dynamic models are not yet known. Future progress in this area will help to ensure that empirical work on the determinants of FDI will be increasingly informative.

<u>STUDY II:</u>

AN INVESTIGATION OF THE HUMAN-CAPITAL AUGMENTED SOLOW MODEL

CHAPTER II.1:

INTRODUCTION

This study aims to estimate the human-capital augmented Solow model (HCASM) which is proposed by Mankiw, Romer and Weil (1992) based on the work by Solow (1956) and Swan (1956). In particular, this study tests the effect of the investment rate, labour-force growth and the level of human capital on the level of the growth path of output per worker, as well as examining the HCASM's prediction of conditional convergence. This study differs from previous cross-country empirical studies partly in that it uses the newest version of the Penn World Table, version 8.0, recently published by Feenstra, Inklaar and Timmer (2013) and the newest Educational Attainment Dataset version 1.3 recently published by Barro and Lee (2013) to estimate the model. To our knowledge, these recent datasets have not yet been used to test the HCASM in the literature.

The second aim of this study is to test whether the HCASM can account for the growth of countries in sub-Saharan Africa and East Asia – the two regions are often known as contrasting regions in terms of economic growth over recent decades.

The structure of this study is as follows. Chapter two presents the literature review. Chapter three discusses the empirical methods and data used in this study. Chapter four reports empirical results and, finally, chapter five provides the conclusion of the research.

CHAPTER II.2:

LITERATURE REVIEW

The literature on economic growth is extensive in terms of both the theory and empirics. Excellent surveys are found in many works such as those by Temple (1999), Barro and Sala-i-Martin (2004), and Durlauf, Johnson and Temple (2005). The growth literature is generally categorized into two main groups being the neoclassical models and the endogenous growth models. The neoclassical models with an exogenous saving rate rely on the studies proposed by Solow (1956) and Swan (1956). Assuming a production function with diminishing returns to capital, the Solow-Swan model (hereafter the Solow model) predicts that the saving rate has a positive effect, while the labour-force growth rate has a negative effect, on a country's steady-state output per worker. Further, it makes a prediction about conditional convergence, namely if structural parameters of countries are similar, those with a lower initial level of output per worker tend to grow faster than those with a higher initial level of output per worker. However, one debateable assumption made in the Solow model is that the long-run rate of growth is determined exogenously by the rate of technological change. By contrast, assuming non-diminishing returns to factors of production, endogenous growth models suggested in studies such as those by Romer (1986) and Lucas (1988) endogenously take the change of technology into account by introducing research and development into the model, treating human capital investment decisions as endogenous, or by assuming a learning-by-doing process. These models often do not support the convergence hypothesis (Mankiw, Romer and Wei, 1992; Barro and Sala-i-Martin, 2004). Although the number of studies based on endogenous growth models increased in the 1980s and 1990s, the neo-classical model is still a popular model in the empirical study of growth. The theoretical framework of the Solow model provides the fundamental specification for many empirical studies, for example those by Mankiw, Romer and Wei (1992) (hereafter MRW), Islam (1995), Caselli, Esquivel and Lefort (1996), Bond, Hoeffler and Temple (2001) and Hoeffler (2002). The studies commonly estimate the Solow model based on a Cobb-Douglass production function. More recently, Duffy and Papageorgiou (2000) and Masanjala and Papageorgiou (2004) extended the Solow model to a constant-elasticity-of-substitution production function. The purpose of this chapter is to focus on the review of the neoclassical model which relies on the Cobb-Douglas production function, which is investigated in our empirical study. This chapter consists of two sections. The first section reviews the basic Solow model, and then the second section reviews the human-capital augmented Solow model (HCASM) and major empirical studies on the HCASM.

The basic Solow model

The basic Solow model[5] assumes a neoclassical production at time *t*:

$$Y_t = F(K_t, A_t L_t) \qquad \text{(II. 1)}$$

where Y_t is output, K_t is physical capital, L_t is labour and A_t is technology. $A_t L_t$ is referred to as effective labour, taking into account labour (L_t) and technology (A_t).

The neoclassical production function has three important assumptions (Barro and Sala-i-Martin, 2004; Romer, 2006). Firstly, it has constant returns to scale in its capital and labour input:

$$F(aK_t, aA_t L_t) = a * F(K_t, A_t L_t) \quad \text{for all } a \geq 0 \qquad \text{(II. 2)}$$

---

[5] The exposition of the basic Solow model in this section relies on Solow (1956), Mankiw, Romer and Weil (1992), Barro and Sala-i-Martin (2004) and Romer (2006).

Under the above assumption, setting $a = \frac{1}{A_t L_t}$ yields the intensive form of the production function

$$\frac{Y_t}{A_t L_t} = F\left(\frac{K_t}{A_t L_t}, 1\right) = \frac{1}{A_t L_t} F(K_t, A_t L_t) \tag{II.3}$$

Define

$$y_t \equiv \frac{Y_t}{A_t L_t}, \; k_t \equiv \frac{K_t}{A_t L_t} \text{ and } f(k_t) \equiv F(k_t, 1) \tag{II.4}$$

where $y_t \equiv \frac{Y_t}{A_t L_t}$ refers to output per effective worker and $k_t \equiv \frac{K_t}{A_t L_t}$ refers to capital per effective worker.

Then, equation (II.3) can be written as

$$y_t = f(k_t) \tag{II.5}$$

The second assumption for the production function is the rule of diminishing returns in capital and labour. This assumption implies that, holding labour and the level of technology constant, the marginal product of capital is positive but it decreases if capital increases. Similarly, under the assumption of diminishing returns to labour, the marginal product of labour is positive but it decreases if labour increases, holding capital and the level of technology constant.

$$\frac{\partial F(K_t, A_t L_t)}{\partial K_t} > 0 \text{ and } \frac{\partial^2 F(K_t, A_t L_t)}{\partial K_t^2} < 0$$

$$\frac{\partial F(K_t, A_t L_t)}{\partial L_t} > 0 \text{ and } \frac{\partial^2 F(K_t, A_t L_t)}{\partial L_t^2} < 0$$

The third assumption is that the production function satisfies the Inada (1963) conditions as follows

$$\lim_{K_t \to 0} \frac{\partial F(K_t, A_t L_t)}{\partial K_t} = \infty \text{ and } \lim_{K_t \to \infty} \frac{\partial F(K_t, A_t L_t)}{\partial K_t} = 0$$

$$\lim_{L_t \to 0} \frac{\partial F(K_t, A_t L_t)}{\partial L_t} = \infty \text{ and } \lim_{L_t \to \infty} \frac{\partial F(K_t, A_t L_t)}{\partial L_t} = 0$$

The assumptions imply that the marginal product of capital (or labour) is very large if capital (or labour) is very small, and the marginal product becomes very small if capital (or labour) is very large.

Let us use the Cobb-Douglas production function as an example to illustrate the basic Solow model

$$Y_t = F(K_t, A_t L_t) = K_t^{\alpha} (A_t L_t)^{1-\alpha} \qquad 0 < \alpha < 1 \tag{II.6}$$

Under the first neoclassical assumption, the Cobb-Douglas production function can be written in the intensive form:

$$y_t = f(k_t) = k_t^{\alpha} \tag{II.7}$$

where $y_t = \frac{Y_t}{A_t L_t}$ and $k_t = \frac{K_t}{A_t L_t}$

It can be seen that equation (II.7) satisfies the second and third neoclassical assumptions.

$$\frac{\partial f(k_t)}{\partial k_t} = \alpha k_t^{\alpha-1} > 0 \text{ and } \frac{\partial^2 f(k_t)}{\partial k_t^2} = \alpha(\alpha - 1) k_t^{\alpha-2} < 0$$

$$\lim_{k \to 0} \frac{\partial f(k_t)}{\partial k_t} = \lim_{k \to 0} (\alpha k_t^{\alpha-1}) = \infty \text{ and } \lim_{k \to \infty} \frac{\partial f(k_t)}{\partial k_t} = \lim_{k \to \infty} (\alpha k_t^{\alpha-1}) = 0$$

We now consider the Solow model in continuous time. Assuming that technology and labour force grow at rates *g* and *n* respectively

$$\frac{\dot{A}_t}{A_t} = g \qquad \text{(II.8)}$$

$$\frac{\dot{L}_t}{L_t} = n \qquad \text{(II.9)}$$

or equivalently

$$A_t = A_0 e^{gt} \qquad \text{(II.10)}$$

$$L_t = L_0 e^{nt} \qquad \text{(II.11)}$$

where $\dot{A}_t$ and $\dot{L}_t$ are derivatives of A_t and L_t with respect to time, g and n are exogenous parameters, and A_0 and L_0 are the values of A and L at $t = 0$.

Then effective labour, $A_t L_t$, grows at rate (n+g).

Suppose that a fixed fraction of output, s, is invested or saved. One unit of output is invested yields one unit of new capital. Additionally, existing capital depreciates at rate δ. Therefore, the change of capital stock can be written as

$$\dot{K}_t = sY_t - \delta K_t \qquad \text{(II.12)}$$

where $\dot{K}_t$ is the derivative of K_t with respect to time.

Dividing both sides of equation (II.12) by $A_t L_t$

$$\frac{\dot{K}_t}{A_t L_t} = s\frac{Y_t}{A_t L_t} - \delta\frac{K_t}{A_t L_t} \qquad \text{(II.13)}$$

or

$$\frac{\dot{K}_t}{A_t L_t} = sy_t - \delta k_t \qquad \text{(II.14)}$$

It can be seen that the left-hand side of equation (II.I4) is the instantaneous change in the capital stock, $\dot{K}_t$, divided by efficiency units of labour, which can be interpreted as net investment per effective worker. The right-hand-side variables are quantities expressed in efficiency units of labour, and it would be useful to write the left-hand-side in terms of the growth rate of capital per effective worker, k_t. This can be done as follows:

Since $k_t = \frac{K_t}{A_t L_t}$, the time-derivative of k_t is

$$\dot{k}_t = \frac{\dot{K}_t A_t L_t - K_t\left(\dot{A}_t L_t + A_t \dot{L}_t\right)}{(A_t L_t)^2} = \frac{\dot{K}_t}{A_t L_t} - \left(\frac{K_t}{A_t L_t}\frac{\dot{A}_t}{A_t} + \frac{K_t}{A_t L_t}\frac{\dot{L}_t}{L_t}\right)$$

$$= \frac{\dot{K}_t}{A_t L_t} - \left(\frac{\dot{A}_t}{A_t} + \frac{\dot{L}_t}{L_t}\right) k_t \qquad (\mathrm{II}.15)$$

Substituting for $\frac{\dot{A}_t}{A_t}$ and $\frac{\dot{L}_t}{L_t}$ from (II.8) and (II.9) respectively into (II.15) gives

$$\dot{k}_t = \frac{\dot{K}_t}{A_t L_t} - (g+n)k_t \qquad (\mathrm{II}.16)$$

Substituting for $\frac{\dot{K}_t}{A_t L_t}$ from (II.14) into (II.16) yields

$$\dot{k}_t = sy_t - (n+g+\delta)k_t \qquad (\mathrm{II}.17)$$

Substituting for y_t from (II.7) into (II.17) yields

$$\dot{k}_t = sf(k_t) - (n+g+\delta)k_t \qquad (\mathrm{II}.18)$$

Equation (II.18) provides an equation of motion for the capital stock per effective worker. It shows that the change in the capital stock per effective worker is equal to the actual investment per effective worker, $sf(k)$, minus the break-even investment per effective worker, $(n+g+\delta)k$. The break-even investment per effective worker can be interpreted as the amount of investment necessary to keep k at the existing level. There are two reasons that some investment is needed to keep k constant: firstly, since existing capital is depreciating, δk is needed to keep the capital stock from falling; secondly, because the quantity of effective labour is growing at rate $n+g$, the $(n+g)k$ is need to provide new capital for the new effective labour to hold k steady (Romer, 2006).

Equation (II.18) implies that k converges to its steady state, k^*. The steady state is determined by setting equation (II.18) equal to zero:

$$\dot{k}_t = 0$$

or

$$sf(k^*) = (n + g + \delta)k^* \tag{II.19}$$

From (II.7) we have $f(k^*) = (k^*)^{\alpha}$. Substituting for $f(k^*)$ into (II.19) yields

$$s(k^*)^{\alpha} = (n + g + \delta)k^* \tag{II.20}$$

or

$$k^* = \left(\frac{s}{n + g + \delta}\right)^{\frac{1}{1-\alpha}} \tag{II.21}$$

From (II.21), we can see the effects of the investment rate and the labour growth rate on output per worker in the steady state as follows. Firstly, take natural logarithms of equation (II.6)

$$\ln Y_t = \alpha \ \ln K_t + (1 - \alpha) \ \ln A_t + (1 - \alpha) \ \ln L_t$$

or

$$\ln\left(\frac{Y_t}{L_t}\right) = \alpha \ \ln\left(\frac{K_t}{L_t}\right) + (1 - \alpha) \ \ln A_t \tag{II.22}$$

We have $k_t = \frac{K_t}{A_t L_t}$ or $\frac{K_t}{L_t} = k_t A_t$ (II.23)

In the steady state: $k_t = k^*$, thus we substitute for k_t from (II.21) into (II.23) which yields

$$\frac{K_t}{L_t} = \left(\frac{s}{n+g+\delta}\right)^{\frac{1}{1-\alpha}} A_t \qquad \text{(II.24)}$$

Substituting for $\frac{K_t}{L_t}$ from (II.24) into (II.22) yields

$$\ln\left(\frac{Y_t}{L_t}\right) = \ln A_t + \frac{\alpha}{1-\alpha}\ln s - \frac{\alpha}{1-\alpha}\ln(n+g+\delta) \qquad \text{(II.25)}$$

Substituting for A_t from (II.10) into (II.25) gives

$$\ln\left(\frac{Y_t}{L_t}\right) = \ln A_0 + gt + \frac{\alpha}{1-\alpha}\ln s - \frac{\alpha}{1-\alpha}\ln(n+g+\delta) \qquad \text{(II.26)}$$

Equation (II.26) predicts that, in the steady state, the investment rate in capital, s, is positively correlated with output per worker while the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation), $n + g + \delta$, is negatively correlated with output per worker. This implies that a country with a higher the investment rate in capital will tend to be richer in terms of per worker output while a country with a higher growth rate of labour (adjusted by the rate of technological progress and the rate of depreciation) will tend to be poorer in per worker output in the steady state.

One concern of the Solow model is the rate at which an economy reaches its steady state. The convergence rate is determined as follows:

Recall equation (II.18) of the Solow model

$$\dot{k}_t = sf(k_t) - (n+g+\delta)k_t$$

Call the right-hand-side of the equation (II.18) as a function $z(k)$. The function $z(k)$ can be rewritten as

$$z(k_t) = z\left(e^{\ln k_t}\right) = sf\left(e^{\ln k_t}\right) - (n + g + \delta)e^{\ln k_t} \qquad \text{(II.27)}$$

Differentiating equation (II.27) with respect to $\ln k_t$ yields

$$\frac{\partial z\left(e^{\ln k_t}\right)}{\partial \ln k_t} = sf'\left(e^{\ln k_t}\right)e^{\ln k_t} - (n + g + \delta)e^{\ln k_t}$$

$$= sf'(k_t)k_t - (n + g + \delta)k_t \qquad \text{(II. 28)}$$

The first-order Taylor series approximation of $z(.)$ with respect to $\ln k_t$ around the steady state $k_t = k^*$ is

$$z(k_t) = z\left(e^{\ln k_t}\right) \approx z\left(e^{\ln k^*}\right) + \left.\frac{\partial z\left(e^{\ln k}\right)}{\partial \ln k}\right|_{k=k^*} (\ln k - \ln k^*) \qquad \text{(II. 29)}$$

In the steady state: $\dot{k}_t = 0$ and hence we have $z(k^*) = z\left(e^{\ln k^*}\right) = 0$. Therefore, when the economy is close to the steady state, equation (II.29) becomes

$$z(k_t) \approx \left.\frac{\partial z\left(e^{\ln k_t}\right)}{\partial \ln k_t}\right|_{k_t=k^*} (\ln k_t - \ln k^*) \qquad \text{(II. 30)}$$

Substituting for $\frac{\partial z\left(e^{\ln k}\right)}{\partial \ln k}$ from (II.28) into (II.30) yields

$$z(k_t) \approx [sf'(k^*)k^* - (n + g + \delta)k^*](\ln k - \ln k^*) \qquad \text{(II. 31)}$$

In addition, in the steady state, from (II.19) we have

$$sf(k^*) = (n\ + g\ +\ \delta)k^*$$

or

$$s = \frac{(n\ + g + \delta)k^*}{f(k^*)} \qquad \text{(II. 32)}$$

Substituting for s from (II.32) into (II.31) gives

$$z(k_t) \approx \left[\frac{(n + g + \delta)k^*}{f(k^*)} f'(k^*)k^* - (n + g + \delta)k^*\right](\ln k - \ln k^*)$$

$$= -\left(1 - \frac{f'(k^*)k^*}{f(k^*)}\right) (n + g + \delta)k^*(\ln k - \ln k^*) \qquad (\text{II}.33)$$

Since $z(k_t) = \dot{k}_t$, thus

$$\dot{k}_t \approx -\left(1 - \frac{f'(k_t)k_t}{f(k_t)}\right)(n + g + \delta)k_t(\ln k_t - \ln k^*) \qquad (\text{II}.34)$$

Since $k_t = \frac{K_t}{A_t L_t}$ and $(k_t) = y_t = \frac{Y_t}{A_t L_t}$, the term $\frac{f'(k)k}{f(k)}$ in (II.34) can be written as

$$\frac{f'(k)k}{f(k)} = \frac{f'(k)\frac{K_t}{A_t L_t}}{\frac{Y_t}{A_t L_t}} = f'(k)\frac{K_t}{Y_t} = r\frac{K_t}{Y_t}$$

where r is the return to capital.

Therefore, this term is the share of capital income in total income, α.

Then, the growth rate of k is

$$\frac{\dot{k}_t}{k_t} \approx -(1 - \alpha)(n + g + \delta)(\ln k_t - \ln k^*) \qquad (\text{II}.35)$$

Define

$$\lambda = (1 - \alpha)(n + g + \delta) \qquad (\text{II}.36)$$

Then, equation (II.35) can be written as

$$\frac{\dot{k}_t}{k_t} \approx (-\lambda)(\ln k_t - \ln k^*) \qquad (\text{II}.37)$$

Equation (II.37) shows how quickly the capital stock per effective worker, k, approaches its steady state, k^*, thus λ is called the rate of convergence. Similarly, we can find that y also converges to its steady state at the same rate as follows: From (II.7), we have

$$y_t = k_t^{\alpha}$$

The above equation implies

$$\frac{\dot{y}_t}{y_t} = \alpha\,\frac{\dot{k}_t}{k_t} \qquad \text{(II. 38)}$$

Substituting for $\frac{\dot{k}_t}{k_t}$ from (II.37) into (II.38), we find that output per effective worker, y, converges to its steady state at rate λ, the same rate as the convergence rate of the capital stock per effective worker, k:

$$\frac{\dot{y}_t}{y_t} = \frac{d\ln y_t}{dt} \approx \alpha(-\lambda)(\ln k_t - \ln k^*) = (-\lambda)(\alpha \ln k_t - \alpha \ln k^*)$$

$$= (-\lambda)(\ln y_t - \ln y^*) \qquad \text{(II. 39)}$$

Equation (II.39) implies that $\ln y_t$ moves toward $\ln y^*$ at a speed approximately proportional to its distance from $\ln y^*$. That is, the growth rate of $\ln y_t - \ln y^*$ is approximately constant and equal to $-\lambda$ (Barro and Sala-i-Martin, 2004, pp.57-58; Romer, 2006, pp.25-26). This implies

$$\ln y_t = \left(1 - e^{-\lambda t}\right) \ln y^* + e^{-\lambda t} \ln y_0 \qquad \text{(II. 40)}$$

where y_0 denotes initial output per effective worker.

Subtracting $\ln y_0$ from both sides of equation (II.40) gives

$$\ln y_t - \ln y_0 = \left(1 - e^{-\lambda t}\right) \ln y^* - \left(1 - e^{-\lambda t}\right) \ln y_0 \qquad \text{(II. 41)}$$

From (II.7), we have $y_t = f(k_t) = k_t^{\alpha}$, thus $y^* = f(k^*) = (k^*)^{\alpha}$ (II. 42)

Substituting for k^* from (II.21) into (II.42) yields $y^* = \left(\frac{s}{n+g+\delta}\right)^{\frac{\alpha}{1-\alpha}}$ (II.43)

Substituting for y^* from (II.43) into (II.41), we have

$$\ln y_t - \ln y_0 = -\rho \ln y_0 + \rho \frac{\alpha}{1-\alpha} \ln(s) - \rho \frac{\alpha}{1-\alpha} \ln(n + g + \delta) \quad (\text{II.} 44)$$

where $\rho = \left(1 - e^{-\lambda t}\right)$.

The above equation expresses that the initial output per effective worker and the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation) are negatively correlated with the growth of output per effective worker, while the investment rate in capital is positively correlated with the growth of output per effective worker. Alternatively, equation (II.44) can be expressed in per worker terms instead of per effective worker terms. In addition, a formulation in per worker terms enables us to test the model empirically. We can reformulate equation (II.44) with respect to output per worker as follows:

We have $y_t = \frac{Y_t}{L_t A_t}$. Substituting for A_t from (II.10) yields

$$y_t = \frac{Y_t}{L_t(A_0 e^{gt})} \quad (\text{II.} 45)$$

Taking logs of both sides of equation (II.45) gives:

$$\ln y_t = \ln\left(\frac{Y_t}{L_t}\right) - \ln A_0 - gt \quad \text{where } \left(\frac{Y_t}{L_t}\right) \text{ is the output per worker} \quad (\text{II.} 46)$$

Substituting for $\ln y_t$ from (II.45) into (II.46) yields:

$$\ln(\frac{Y_t}{L_t}) - \ln(\frac{Y_0}{L_0}) = -\rho \ln(\frac{Y_0}{L_0}) + \rho \frac{\alpha}{1-\alpha} \ln(s) - \rho \frac{\alpha}{1-\alpha} \ln(n + g + \delta)$$

$$+\rho \ln A_0 + gt \qquad \text{where } \rho = \left(1 - e^{-\lambda\tau}\right) \qquad (\text{II}.47)$$

Equation (II.47) implies that in the basic Solow model the growth rate of per worker output depends on initial output per worker, $\frac{Y_0}{L_0}$, the investment rate in capital, s, the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation), $n + g + \delta$, the initial level of technology, A_0, and the rate of technological progress, g. Therefore, the basic Solow model predicts that initial output per worker and the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation) are negatively correlated with the growth of output per worker while the saving rate is positively correlated with the growth of output per worker. In addition, equation (II.47) indicates that the coefficient on the investment rate in capital is $\left(\rho \frac{\alpha}{1-\alpha}\right)$ while that on the growth of the labour force (adjusted by the rate of technological progress and the rate of depreciation) is $\left(-\rho \frac{\alpha}{1-\alpha}\right)$, and thus the basic Solow model predicts that the coefficients on the investment rate in capital and the growth of the labour (adjusted by the rate of technological progress and the rate of depreciation) have the same magnitude and opposite signs.

The human-capital augmented model (HCASM)

MRW argue that economists have long emphasized the importance of human capital to growth, and thus they suggest augmenting the basic Solow model by introducing human capital into the Solow model in order to consider the role of human capital in the process of growth. The Cobb-Douglas production function is then given by

$$Y_t = K_t^{\alpha} H_t^{\beta} (A_t L_t)^{1-\alpha-\beta} \qquad (\alpha + \beta) < 1 \qquad (\text{II}.48)$$

where α and β are the exponents on physical and human capital respectively.

MRW assumed that there are diminishing returns to capital as a whole, and thus the sum of α and β is less than one. Denote s_k and s_h as the proportions of output invested in physical and human capital respectively.

Assuming that physical capital and human capital depreciate at the same rate[6], δ, and using the same steps as in the previous section, the increase in physical capital per effective worker and human capital per effective worker at a point in time is

$$\dot{k}_t = s_k y_t - (n + g + \delta)k_t = s_k k_t^{\alpha} h_t^{\beta} - (n + g + \delta)k_t \qquad \text{(II.49)}$$

$$\dot{h}_t = s_h y_t - (n + g + \delta)h_t = s_h k_t^{\alpha} h_t^{\beta} - (n + g + \delta)h_t \qquad \text{(II.50)}$$

where

k_t is defined as the stock of physical capital per effective worker, $k_t = \frac{K_t}{A_t L_t}$

h_t is defined as the stock of human capital per effective worker, $h_t = \frac{H_t}{A_t L_t}$,

and y_t is defined as the level of output per effective worker, $y_t = \frac{Y_t}{A_t L_t}$.

Equation (II.49) and equation (II.50) imply that the physical capital per effective worker, k_t, and the human capital per effective worker, h_t, converge to their steady states, $k^*$and h^*, respectively, where

$$\dot{k}_t = 0$$

$$\dot{h}_t = 0$$

---

[6] This assumption could be restrictive. However, assuming different depreciation rates for physical capital and human capital is unlikely to make a great deal of difference in the MRW approach: as an empirical matter, it leads to a model with two labour-force growth terms which are almost perfectly correlated (Temple, 1998, p.42).

or

$$s_k(k^*)^{\alpha}(h^*)^{\beta} = (n+g+\delta)k^* \qquad (\text{II}.51)$$

$$s_h(k^*)^{\alpha}(h^*)^{\beta} = (n+g+\delta)h^* \qquad (\text{II}.52)$$

or

$$(k^*)^{\alpha}(h^*)^{\beta} = \frac{(n+g+\delta)k^*}{s_k} \qquad (\text{II}.53)$$

$$(k^*)^{\alpha}(h^*)^{\beta} = \frac{(n+g+\delta)h^*}{s_h} \qquad (\text{II}.54)$$

From equations (II.53) and (II.54), we have

$$\frac{k^*}{s_k} = \frac{h^*}{s_h}$$

or

$$k^* = h^* \frac{s_k}{s_h} \qquad (\text{II}.55)$$

$$h^* = k^* \frac{s_h}{s_k} \qquad (\text{II}.56)$$

Substituting for h^* and k^* from (II.55) and (II.56) into equations (II.51) and (II.52) respectively gives

$$k^* = \left[\frac{\left(s_k^{1-\beta} s_h^{\beta}\right)}{(n+g+\delta)}\right]^{1/(1-\alpha-\beta)} \qquad (\text{II}.57)$$

$$h^* = \left[\frac{(s_k^{\alpha} s_h^{1-\alpha})}{(n+g+\delta)}\right]^{1/(1-\alpha-\beta)} \qquad (\text{II}.58)$$

The steady state output per worker can be determined by substituting equations (II.57) and (II.58) into the production function (II.48) and then taking logs

$$\ln(\frac{Y_t}{L_t}) = \ln A_0 + gt + \frac{\alpha}{1-\alpha-\beta}\ln(s_k) + \frac{\beta}{1-\alpha-\beta}\ln(s_h)$$
$$-\frac{\alpha+\beta}{1-\alpha-\beta}\ln(n+g+\delta) \qquad (\text{II}.59)$$

Equation (II.59) predicts that, in the steady state, the investment rate in physical capital, s_k, and the investment rate in human capital, s_h, are positively correlated with output per worker while the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation), $n + g + \delta$, is negatively correlated with output per worker. This implies that countries with higher investment rates in physical capital and human capital will tend to be richer while countries with a higher growth rate of the labour force will tend to be poorer in terms of output per worker in the steady state. MRW also present another way to consider the role of the human capital in the process of growth by using the level of human capital, h^*, instead of the rate of investment in the human capital, s_h ,as follows.

Taking logs of both sides of equation (II. 58) yields

$$\ln h^* = \frac{\alpha}{1-\alpha-\beta}\ln(s_k) + \frac{1-\alpha}{1-\alpha-\beta}\ln(s_h) - \frac{1}{1-\alpha-\beta}\ln(n+g+\delta)$$

or

$$\frac{1}{1-\alpha-\beta}\ln(s_h) = \frac{1}{1-\alpha}\ln h^* - \frac{1}{1-\alpha}\frac{\alpha}{1-\alpha-\beta}\ln(s_k)$$
$$+\frac{1}{1-\alpha}\frac{1}{1-\alpha-\beta}\ln(n+g+\delta) \qquad (\text{II}.60)$$

Multiply both sides of equation (II.60) with β to yield

$$\frac{\beta}{1-\alpha-\beta}\ln(s_h) = \frac{\beta}{1-\alpha}\ln h^* - \frac{\beta}{1-\alpha}\frac{\alpha}{1-\alpha-\beta}\ln(s_k)$$

$$+\frac{\beta}{1-\alpha}\frac{1}{1-\alpha-\beta}\ln(n+g+\delta) \qquad \text{(II.61)}$$

Substituting for the term $[\frac{\beta}{1-\alpha-\beta}\ln(s_h)]$ from (II.61) into (II. 59) gives

$$\ln(\frac{Y_t}{L_t}) = \ln A_0 + gt + \frac{\alpha}{(1-\alpha)}\ln(s_k) + \frac{\beta}{1-\alpha}\ln(h^*)$$

$$-\frac{\alpha}{(1-\alpha)}\ln(n+g+\delta) \qquad \text{(II.62)}$$

Equation (II.62) implies that a country with a higher investment rate in physical capital and a higher level of human capital will tend to be richer while a country with a higher growth rate of the labour force will tend to be poorer in terms of output per worker in the steady state.

Similarly as the previous section on the basic Solow model, in order to take into account transitional dynamics for equation (II.62), we call y^* the steady state level of outcome per effective worker and y_t as its actual value at time *t*. Approximating around the steady state, the convergence speed in the HCASM is computed by

$$\frac{d\ln \mathrm{y_t}}{dt} = \lambda[\ln y^* - \ln y_t] \quad \text{where } \lambda = (n+g+\delta)(1-\alpha-\beta) \qquad \text{(II.63)}$$

The equation (II.63) implies

$$\ln y_t = \left(1-e^{-\lambda t}\right)\ln y^* + e^{-\lambda t}\ln y_0 \qquad \text{(II.64)}$$

where $y_0$is the initial output per effective worker.

Subtracting $\ln y_0$from both sides of the equation yields

$$\ln y_t - \ln y_0 = \left(1-e^{-\lambda t}\right)\ln y^* - \left(1-e^{-\lambda t}\right)\ln y_0 \qquad \text{(II.65)}$$

We now try to compute y^* as follows. Firstly, multiplying both sides of (II.48) by $\left(\frac{1}{A_t\,L_t}\right)$ yields

$$\frac{Y_t}{A_tL_t}=\left(\frac{K_t}{A_tL_t}\right)^{\alpha}\left(\frac{H_t}{A_tL_t}\right)^{\beta}\ \text{ or }\ y_t=k_t^{\alpha}h_t^{\beta} \qquad \text{(II. 66)}$$

From (II.66), y^* is given by

$$y^*=(k^*)^{\alpha}(h^*)^{\beta} \qquad \text{(II. 67)}$$

Substituting for k^* and k^* from (II. 57) and (II. 58) respectively into (II. 67) gives

$$y^*=\left[\frac{\left(s_k^{1-\beta}s_h^{\beta}\right)}{(n+g+\delta)}\right]^{\alpha/(1-\alpha-\beta)} * \left[\frac{(s_k^{\alpha}s_h^{1-\alpha})}{(n+g+\delta)}\right]^{\beta/(1-\alpha-\beta)} \qquad \text{(II. 68)}$$

Substituting for y^* from (II. 68) into (II.65) yields

$$\ln y_t-\ln y_0=-\rho\ln y_0+\rho\frac{\alpha}{1\text{–}\alpha-\beta}\ln(s_k)+\rho\frac{\beta}{1\text{–}\alpha-\beta}\ln(s_h)$$

$$-\rho\frac{\alpha+\beta}{1\text{–}\alpha-\beta}\ln(n+g+\delta)\ \text{ where } \rho=\left(1-e^{-\lambda t}\right) \qquad \text{(II. 69)}$$

Equation (II.69) expresses the correlation between the growth of output per effective worker and the initial output per effective worker, y_0, the investment rate in physical capital, s_k, the investment rate in human capital, s_h, and the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation), $n+g+\delta$. Alternatively, equation (II.69) can be expressed in per worker terms instead of in per effective worker terms. In addition, a formulation in per worker terms enables us to test the model empirically. We can reformulate equation (II.44) with respect to output per worker as follows:

We have

$$y_t = \frac{Y_t}{L_t A_t} = \frac{Y_t}{L_t(A_0 e^{gt})} \tag{II.70}$$

Taking logs both side of equation (II.70) gives:

$$\ln y_t = \ln\left(\frac{Y_t}{L_t}\right) - \ln A_0 - gt \quad \text{where } \left(\frac{Y_t}{L_t}\right) \text{ is output per worker} \tag{II.71}$$

Substituting for $\ln y_t$ from (B.75) into equation (B.74) yields

$$\ln\left(\frac{Y_t}{L_t}\right) - \ln\left(\frac{Y_0}{L_0}\right) = -\rho \ln\left(\frac{Y_0}{L_0}\right) + \rho\frac{\alpha}{1-\alpha-\beta}\ln(s_k) + \rho\frac{\beta}{1-\alpha-\beta}\ln(s_h)$$

$$-\rho\frac{\alpha+\beta}{1-\alpha-\beta}\ln(n+g+\delta) + \rho\ln A_0 + gt \tag{II.72}$$

where $\rho = \left(1 - e^{-\lambda t}\right)$ and λ is the convergence rate to the steady state.

Equation (II.72) now represents the relationship between the growth of output per worker and the initial level of output per worker, $\frac{Y_0}{L_0}$, the investment rate in physical capital, s_k, the investment rate in human capital, s_h, the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation), $n+g+\delta$, the initial level of technology, A_0, and the rate of technological progress, g, in the HCASM. The model predicts that the initial level of output per worker and the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation) negatively impact the growth rate of output per worker while the investment rate in physical capital and the investment rate in human capital positively impact the growth rate of output per worker. In addition, equation (II.72) implies that the coefficients on the investment rate in physical capital and the investment rate in human capital are $\left(\rho\frac{\alpha}{1-\alpha-\beta}\right)$ and $\left(\rho\frac{\beta}{1-\alpha-\beta}\right)$ respectively while that on the growth of the labour force (adjusted by the rate of technological progress and the rate of depreciation) is $-\left(\rho\frac{\alpha+\beta}{1-\alpha-\beta}\right)$, and thus the HCASM further predicts that the sum of the coefficients on the investment rate in physical capital, the investment rate in human capital and the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation) is zero.

Alternatively, the effect of human capital on the growth of output per worker can be analysed by using the level of human capital (instead of the investment rate in human capital). To consider the effect of the level of the human capital, we can substitute for the term $\left[\frac{\beta}{1-\alpha-\beta}\ln(s_h)\right]$ from (II.60) into equation (II.72). This leads to

$$\ln\left(\frac{Y_t}{L_t}\right)-\ln\left(\frac{Y_0}{L_0}\right)=-\rho\ln\left(\frac{Y_0}{L_0}\right)+\rho\frac{\alpha}{1-\alpha}\ln(s_k)+\rho\frac{\beta}{1-\alpha}\ln(h)$$

$$-\rho\frac{\alpha}{1-\alpha}\ln(n+g+\delta)+\rho\ln A_0+gt \qquad (\text{II}.73)$$

where $\rho=\left(1-e^{-\lambda t}\right)$ and λ is the convergence rate to the steady state.

In this study, the HCASM described by equation (II.73) is used in the empirical analysis. Equation (II.73) implies that in the HCASM the growth rate of output per worker depends on initial output per worker, $\frac{Y_0}{L_0}$, the investment rate in physical capital, s_k, the level of human capital, h, the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation), $n+g+\delta$, the initial level of technology, A_0, and the rate of technological progress, g. The model predicts that initial output per worker and the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation) are negatively correlated with the growth of output per worker while the investment rate in physical capital and the level of human capital are positively correlated with the growth of output per worker. In addition, equation (II.47) implies that the coefficient on the investment rate in physical capital is $\left(\rho\frac{\alpha}{1-\alpha}\right)$ while that on the growth of the labour force (adjusted by the rate of technological progress and the rate of depreciation) is $\left(-\rho\frac{\alpha}{1-\alpha}\right)$, and thus the HCASM predicts that the coefficients on the investment rate in physical capital and the growth of labour (adjusted by the rate of technological progress and the rate of depreciation) have the same magnitude and opposite signs. Equations (II.72) or (II.73) are used as a framework for empirical research in MRW, Islam (1995), Caselli, Esquivel and Lefort (1996), Bond, Hoeffler and Temple (2001) and Hoeffler (2002).

Using a cross-sectional sample including 98 countries over the period 1960-1985, MRW test the HCASM by using an OLS estimator. They (p.410) argue that g reflects the advancement of knowledge which is not country-specific. And there is neither any strong reason to expect δ to vary greatly across countries, nor are there any reliable data that would allow us to estimate country-specific depreciation rates, and thus MRW assume the rate of technological progress, g, and the depreciation rate of capital, δ, to be constant across countries (MRW, p.410). However, they allow the initial level of technology, A_0, to be different across countries:

$$\ln(A_0)_i = c + \epsilon_i$$

where c is a constant, ϵ is a country-specific shock term, and the subscript i denotes the country.

In the MRW, the investment rate in physical capital and the growth of the labour force are proxied by the average investment share in GDP and the average growth rate of the working-age population respectively while the investment rate in human capital is proxied by the secondary school enrolment rate of the population multiplied by the fraction of the working-age population that is of school age. Assuming that the investment rate in capital and the growth rate of labour are uncorrelated with the error term, ϵ, they performed OLS to test the HCASM that has the form in $(\mathrm{II}.72)$.

Empirical results in MRW support the HCASM: the coefficient on initial output per worker (i.e. the output per worker in 1960) is found to be significant and negative, implying evidence of conditional convergence. The conditional convergence implies that a country with a lower initial output per worker tends to grow faster than another country with higher initial output per workers once the other regressors including the saving rate, human capital and the growth of the labour force are held constant.[7] The rate of the convergence[8], λ, is found to

---

[7] Note that according to the conditional convergence in the (human-capital augmented) Solow model a poor country may grow slower than a rich country does (MRW; Barro and Sala-i-Martin, 2004). For example, suppose that there are two countries with differences in initial output per worker and investment rates in capital. If the rich country (i.e. the country with higher initial output per worker) has a higher investment rate in capital than the poor country does, the rich country would have a higher output per worker in the steady state, and thus the rich country may be proportionally further from its steady state. In this case, the growth of the output per worker of the rich country may be higher than that of the poor country.

[8] In the empirics of growth, a finding of a significantly negative relationship between the initial output per worker and the growth of the output per worker is evidence for conditional convergence (MRW; Durlauf, Johnson and Temple, 2005). The rate of convergence is calculated from the estimated coefficient on the initial output per worker as follows.

From equation (II.72) [or equation (II.73)], we have that the coefficient on the initial output per worker in the human-capital augmented Solow model is $-\left(1-e^{-\lambda t}\right)$ where λ is the rate of convergence and *t* is the length of the time period.

Set that the estimated coefficient on the initial output per worker in an empirical estimation of the human-capital augmented Solow model is η, and then:

$$\eta = -\left(1-e^{-\lambda t}\right)$$

Therefore, the rate of convergence is computed as

$$\lambda = -\frac{\ln(\eta+1)}{t}$$

be about 1.4 percent a year.[9] In addition, the coefficients on the investment rate in physical capital, the investment rate in human capital and the growth of the labour force are significant with the expected signs, implying that the investment rates in the physical and human capital positively impact the growth of output per worker while the growth of the labour force negatively impacts the growth of output per worker. Also, the restriction that the sum of the coefficients on $\ln(s_k)$, $\ln(s_h)$ and $\ln(n+g+\delta)$ is equal to zero is not rejected.

Islam (1995) shows that the assumption that the investment rate in capital and the growth of the labour force are independent of the country-specific shocks, ϵ, in the cross-sectional regression using OLS in MRW may be too strong. This is because ϵ may reflect not only technology but also other factors such as resource endowment and institutions, and thus it may be potentially correlated with capital accumulation and the growth rate of the labour force. This potentially leads to bias in the estimates in the OLS estimation. Islam suggests using panel data by replacing MRW's single cross-section over the entire period with cross sections for shorter periods, and using a fixed effects (FE) estimator to control for country-specific intercepts ("fixed effects"). Using a sample including 79 countries over five-year intervals for the period 1960-1985, he tests the HCASM that has the form in (II.73) where the level of human capital is proxied by the average schooling years in the total population

---

[9] The convergence rate, λ, being 1.4 percent implies that the average time an economy takes to cover half of the distance between its initial position to its steady state (hereafter the half-life) is about 50 years.
Generally, the half-life is calculated as follows: Recall equation (II.65): $\ln y_t - \ln y_0 = \left(1-e^{-\lambda t}\right)\ln y^* - \left(1-e^{-\lambda t}\right)\ln y_0$. This equation leads: $\ln y_t = \left(1-e^{-\lambda t}\right)\ln y^* + e^{-\lambda t}\ln y_0$. The half-life, t, for which $\ln y_t$ is halfway between $\ln y_0$ and $\ln y^*$satisfies the condition $e^{-\lambda t} = 1/2$. Taking logs both sides yields: $t = -\ln(1/2)/\lambda = 0.69/\lambda$ (Caselli, Esquivel and Lefort, 1996; Barro and Sala-i-Martin, 2004; Romer, 2005). Therefore, if $\lambda = 1.4$ percent per year, the half-life is about 50 years.

over age 25. Empirical results show the coefficients on the initial output per worker and population growth are significant and negative while that on the investment rate in physical capital is significant and positive. The rate of convergence to the steady state is found to be approximately 3.8 percent a year. However, the coefficient on the human capital is found to be significant but *negative*. Therefore, Islam concludes that the empirical results do not provide much support for the HCASM.

Caselli, Esquivel and Lefort (1996) (hereafter CEL) show that the FE estimation may give biased estimates because the HCASM estimated is a dynamic model and *T*, the number of time periods, is small. In addition, the standard FE estimator cannot address the potential problem of endogeneity if one or more variables in the right-hand side of the growth models are endogenous. In order to address these problems, CEL suggest using Arellano and Bond (1991)'s first-differenced generalised method of moments (DGMM) estimator that uses the second and earlier lagged values of endogenous variables as instruments for the subsequent first-differences of the endogenous variables to test the HCASM. Using panel data at five-year intervals in a sample including 97 countries for 1960-1985, CEL estimate the HCASM that has the form in (II.72) where the investment rate in human capital is proxied by the secondary school enrolment rate (different from MRW). Empirical results in CEL show that the coefficient on initial output per worker is significant and negative, indicating the existence of conditional convergence. The convergence rate is found to be 6.8 percent a year. In addition, the coefficients on the investment rate in physical capital and the growth of the labour force are significant with the right signs as the HCASM predicts. However, the coefficient on the investment rate in human capital is found to be significant and negative, implying that human capital negatively impacts the growth of output per worker. Therefore, CEL reject the HCASM.

However, Bond, Hoeffler and Temple (2001) (hereafter BHT), based on the work by Blundell and Bond (1998), point out that the lagged levels of persistent variables are weakly correlated with their subsequent first-differences, and thus the instruments for the first-differences in the DGMM estimation of the HCASM in CEL are potentially weak. This may lead to bias for the estimates in CEL. Therefore, in order to deal with the problem of weak instruments in the DGMM estimator, BHT suggest using Blundell and Bond (1998)'s system generalised method of moments (SGMM) estimator that uses lagged first-differences of variables as instruments for equations in levels in addition to lagged levels of the endogenous variables as instruments for equations in first differences. Using the same sample as in Caselli, Esquivel and Lefort (1996), BHT perform a one-step SGMM estimation for the HCASM. Empirical results in BHT show that the coefficients on initial output per worker, the investment rate in physical capital and the growth of the labour force are significant with the expected signs, implying that the investment rate in physical capital positively impacts the growth of output per worker while the initial output per worker and the growth of the labour force negatively impact the growth of output per worker. The implied rate of convergence is 1.7 percent a year. The coefficient on the investment rate in human capital is found to be insignificant.

Parallel to the above works, there are a number of empirical cross-country studies that extend the Solow growth model by introducing additional variables. Most of these works do not explicitly put forward a theoretical framework. Many of the models in these papers are constructed on an *ad hoc* basis. Typically these models embrace initial output per capita, investment rate and additional variables that lie outside the Solow model such as human capital, and measures of policy outcomes in order to investigate different growth determinants or to explain the growth for a certain specific region; and the choice of variables to include varies greatly among works (Hoeffler, 1998; Temple, 1999; Durlauf, Johnson and Temple, 2005). For example, Barro (1991) uses a sample of 98 countries for 1960-1985 to investigate effects of initial GDP per capita, school enrolment, government expenditure, public investment, fertility, political instability and market distortions on the growth of GDP per capita. He includes dummies for sub-Saharan Africa and Latin America in growth models and finds that the coefficients on these two regional dummies are significant and negative in cross-sectional estimation results. Therefore, Barro (1991) concludes that the growth models in those papers do not fully explain the differences in growth between sub-Saharan Africa and Latin America and other regions and admits that some regularity may be missing from his regression model.

Another study by Levine and Renelt (1992) also extends normal growth models in earlier empirical growth studies, by introducing additional variables including government expenditure, economic policies, and political factors into the model, in many possible combinations. Dummies for countries in sub-Saharan Africa and Latin America are also added into their regression models. The sample used in Levine and Renelt (1992) includes 119 countries for 1960-1989 with two datasets: one from the World Bank and IMF and the other from Summers and Heston (1988). The cross-sectional regression results in the research by Levine and Renelt (1992) show that only a few explanatory variables are robustly correlated with growth. In particular, only initial GDP per capita and the investment rate are robust in their regressions.

Subsequent studies, for example those by Sachs and Warner (1997) and Easterly and Levine (1998), have focused on the slow growth of countries in the sub-Saharan African region. Sachs and Warner (1997) try to explain the abnormal growth performance of sub-Saharan African countries by introducing additional variables omitted in previous research into the neoclassical growth model. Using OLS to perform a cross-section regression for the growth model using a sample of 83 countries over the period 1965-1990, Sachs and Warner (1997) indicate that the lack of international trade openness, low life expectancy, and geographical factors, for example tropical climate and being landlocked, may cause the slow growth of sub-Saharan Africa countries.

Another study by Easterly and Levine (1998) uses a growth model including variables for initial GDP per capita, human capital, political instability, financial development, the black market exchange rate premium and the government surplus to GDP ratio to explain the slow growth in sub-Saharan African countries. Using a cross-sectional regression on the model for a sample of countries over the period from 1960 to 1990, Easterly and Levine (1998) find that the coefficient on the sub-Saharan African dummy is significant and negative. However, this dummy is found to be insignificant once the neighbours' growth rate is introduced into the regression model. Therefore, they argue that countries in a region that all have poor policies would each have poor growth performance not only due to their own poor policies, but also due to poor policies of their neighbours.

However, Hoeffler (2002) shows that once unobserved country-specific effects and endogeneity are controlled for, the sub-Saharan African dummy is insignificant even in the HCASM. She argues that the significance of the sub-Saharan African dummy is due to problems related to estimation methods, because earlier studies investigating the poor growth performance of countries in sub-Saharan Africa have not taken account of unobserved country-specific effects. In addition, the earlier studies have not controlled for the potential problem of endogeneity. Using a sample comprising 85 countries for the period 1960-1990, Hoeffler (2002) performs a one-step SGMM estimator that is suggested by BHT to control for unobserved country-specific effects and endogeneity in the HCASM.[10] Then, she regresses the residuals which are obtained from the one-step SGMM estimation on the dummy for sub-Saharan African countries. The estimation results show that the coefficient on sub-Saharan African dummy is insignificant. Therefore, Hoeffler concludes that the HCASM can account for the slower growth of sub-Saharan African countries.

In sum, this chapter provides a theoretical basis for the (human-capital augmented) Solow model. The major empirical studies on the human-capital augmented Solow model are also reviewed in the latter sections of this chapter. In the next chapter, we will discuss the methodology used to test the human-capital augmented Solow model in this study.

---

[10] The one-step system GMM estimation of the HCASM in Hoeffler (2002) also indicates conditional convergence with the rate of convergence being 3.3 percent a year.

CHAPTER II.3:

METHODOLOGY

This chapter consists of the following three sections. The first one discusses the estimation methods used to test the human-capital augmented Solow model (HCASM), and then the second one presents the method used to check whether the HCASM can explain the abnormal growth in sub-Saharan Africa and East Asia. The third section describes the samples and data sources used in this study.

II.3.1. Estimating the HCASM

In the early 1990s, empirical works on economic growth, typically similar to that by MRW, use a *single* cross-section estimation to analyse the HCASM. Such works often use average data for 25 or 30 years. In their single cross-section analysis, the regression model has the following form

$$\Delta y_i = c + \eta y_{i,0} + \theta' X_i + \epsilon_i \qquad \text{for } i = 1, \ldots, N \qquad (\text{II}.74)$$

where subscript i denotes a country index, c is the constant and ϵ_i is the error term; Δy_i is the difference in the natural logarithm of GDP per worker over the entire period observed, e.g. in MRW, Δy_i is the difference in the natural logarithm of GDP per worker over the period 1960-1985, $\Delta y_i = y_{i,1985} - y_{i,1960}$; $y_{i,0}$ is the natural logarithm of GDP per worker at the start of the period observed, e.g. in MRW, $y_{i,0}$ is the natural logarithm of GDP per worker in 1960; and X_i is a vector of variables including the investment rate in physical capital, investment in human capital, the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation).

However, the single cross-section estimation will be not used in this study for the following reasons. Firstly, the use of a single cross-section means not all available information is used (Hoeffler, 2002). It is clear that panel data can utilise more information than a single cross-section of data (Baltagi, 2008). In addition, in the empirical literature on the (human-capital augmented) Solow model, the rate of technological progress is typically assumed to be constant within a period; this assumption may be more realistic over shorter periods of time, e.g. five-year periods, in panel data analysis than over the entire period in a single cross-section analysis (Islam, 1995). Further, a single cross-section estimator cannot control for country-specific time-invariant effects and also makes it harder to address the potential endogeneity of regressors (CEL; Temple, 1999; BHT). These problems can be addressed in the dynamic panel data analysis which is discussed below.

Consider the HCASM from equation (II.73)

$$\ln\left(\frac{Y_t}{L_t}\right) - \ln\left(\frac{Y_0}{L_0}\right) = -\rho\ln\left(\frac{Y_0}{L_0}\right) + \rho\frac{\alpha}{1-\alpha}\ln(s_k) + \rho\frac{\beta}{1-\alpha}\ln(h)$$

$$-\rho\frac{\alpha}{1-\alpha}\ln(n+g+\delta) + \rho\ln A_0 + gt \qquad (\text{II}.73)$$

where $\rho = \left(1 - e^{-\lambda t}\right)$.

In order to test the HCASM, an alternative approach is to use panel data by splitting the single cross-section over the entire period into five-year intervals (Islam, 1995; CEL; BHT). The regression model in the panel estimation has the following generic form:

$$\Delta y_{it} = \eta y_{i,t-1} + \theta' X_{it} + \mu_i + \omega_t + v_{it} \text{ for } i = 1, \ldots, N \text{ and } t = 2, \ldots, T \quad (\text{II}.75)$$

where subscript i denotes a country index and subscript t denotes an interval index; $\Delta y_{it} = y_{it} - y_{i,t-1}$ is the difference in the natural logarithm of GDP per worker over a five-year interval, $y_{i,t-1}$ and y_{it} are the natural logarithm of

GDP per worker at the start and at the end of that interval respectively; μ_i denotes unobserved country-specific effects that partly reflect the difference in the initial level of technology; ω_t denotes the time effects capturing the impacts of technological progress and shocks that are common to all countries (Islam, 1995); X_{it} is a vector of variables including the investment rate in physical capital, human capital, the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation); η and θ are the coefficients on $y_{i,t-1}$ and X_{it} respectively, and v_{it} are the transient errors.

Equation (II. 75) can be equivalently re-written as the following dynamic panel data model

$$y_{it} = \kappa y_{i,t-1} + \theta' X_{it} + \mu_i + \omega_t + v_{it} \text{ for } i = 1, \dots, N \text{ and } t = 2, \dots, T \quad \text{(II. 76)}$$

where $\kappa = \eta + 1$.

The model (II.76) is used as a regression model to test the HCASM in this study.[11] It can be seen that in the Pooled OLS (POLS) estimator that uses a conventional least squares regression, pooling of all (five-year) observations

---

[11] Based on the estimation of the regression model (II.75), to check whether the conditional convergence exists and calculate the convergence rate see footnote 8 on pp.73-74. In the case based on the estimation of the regression model (II.76) as in this study, if the coefficient on initial output per worker is found to be significant and less than one, there exists conditional convergence as the HCASM predicts. The rate of convergence is calculated from the estimated coefficient on the initial output per worker, κ, as follows.
We have $\kappa = \eta + 1$, leading $\eta = \kappa - 1$. From footnote 8, we have

$$\eta = -\left(1 - e^{-\lambda t}\right)$$

or

$$\kappa - 1 = -\left(1 - e^{-\lambda t}\right)$$

Therefore, the rate of convergence is calculated as

$$\lambda = -\frac{\ln(\kappa)}{t}$$

We use five-year-interval data, thus t=5. This yields

$$\lambda = -\frac{\ln(\kappa)}{5}$$

Note that if κ is found to be significant and larger than one, then conditional convergence (as the HCASM predicts) is absent.

without considering country-specific effects (μ_i) could lead to biased estimates. This is because y_{it} is a function of μ_i, thus $y_{i,t-1}$ is also a function of μ_i, and the omission of μ_i leads the error term to become $(\mu_i + v_{it})$ and hence to be positively correlated with (at least) $y_{i,t-1}$. Therefore, the POLS estimate of the coefficient on the lagged dependent variable, $\hat{\kappa}$, is biased upwards (for example, Hsiao, 1986).

The fixed effects estimation method can control for country-specific effects, μ_i , by using a Within Groups (WG) estimator that transforms the equation (II.76) in order to eliminate μ_i. In the WG estimator, the average values of terms in the equation over time for each country i are obtained, and the original observations are expressed as deviations from these averages; then OLS is used to estimate the transformed equation (for example, Bond, 2002; Baltagi, 2008). Because μ_i are invariant over time, these country-specific effects are eliminated from the transformed equation. However, this transformation process leads to a correlation between the transformed lagged dependent variable and the transformed error term. Particularly, in the WG estimator, the transformed lagged dependent variable is $y^*_{i,t-1} = y_{i,t-1} - \frac{1}{T-1}(y_{i1} + \cdots + y_{it} + \cdots + y_{i,T-1})$ whereas the transformed error term is $v^*_{it} = v_{it} - \frac{1}{T-1}(v_{i2} + \cdots + v_{i,t-1} + \cdots + y_{iT})$. The problem is that $y_{i,t-1}$ in $y^*_{i,t-1}$ correlates with $-\frac{1}{T-1}v_{i,t-1}$ in v^*_{it} while $-\frac{1}{T-1}y_{it}$ in $y^*_{i,t-1}$ also correlates with v_{it} in v^*_{it}. This leads to a bias downwards in the WG estimate of the coefficient on the lagged dependent variable, $\hat{\kappa}$ (Nickell, 1981; Bond, 2002; Baltagi, 2008). This bias is firstly found by Nickell (1981) who shows that the WG estimator is biased of O(1/T). The Nickell (1981) bias does not vanish as the number of countries, *N*, in the sample increases, but it will tend to diminish as T increases. The estimates in growth regressions using WG estimation are likely to be biased substantially because the samples employed often have large *N* but only small *T*, as in this study.

Kiviet (1995, 1999), Judson and Owen (1999) and Bun and Kiviet (2003) suggest using a bias-corrected least square dummy variable (LSDVC) estimator to correct for the Nickell (1981) bias in the WG (or least squares dummy variable, LSDV) estimator for balanced panel data samples with large *N* and small *T*.[12] Later, Bruno (2005) extended the LSDVC estimator to unbalanced panels. The approximation to the bias that is used to make the correction requires a consistent estimator, which is often a first-differenced generalised method of moments estimator or a system generalised method of moments estimator, to start from. Since the first-differenced generalised method of moments estimator may suffer a bias due to weak instruments (Blundell and Bond, 1998; Bruno, 2005; see also the discussion on first-differenced and system generalised method of moments estimators below), the system generalised method of moments estimator is used to initialize the bias correction in LSDVC estimation in this study. The standard errors are bootstrapped based on 1000 replications (for example, Kiviet, 1995 and Judson and Owen, 1999). However, the LSDVC estimator, as well as the POLS and WG estimators, is based on an assumption that the explanatory variables (other than the lagged dependent variable) are strictly exogenous. Therefore, they cannot address the potential endogeneity of the explanatory variables. The problem can be addressed by using Arellano and Bond (1991)'s first-differenced generalised method of moments (DGMM) estimator as follows.

---

[12] Kiviet (1995, 1999), Judson and Owen (1999), and Bun and Kiviet (2003) also suggest approaches to correct bias in the WG or LSDV estimator for samples with small *N* or with moderate *N* and *T*.

Consider the model (II.76) with X_{it} including $(k-1)$ explanatory variables

$$y_{it} = \kappa y_{i,t-1} + \theta' X_{it} + \mu_i + \omega_t + v_{it} = \delta' W_{it} + \mu_i + \omega_t + v_{it} \qquad \text{(II.77)}$$

where $W_{it} = \left(y_{i,t-1}\, X_{it}\right)'$ is $k \times 1$.

The DGMM approach firstly differences the model (II.77) to remove the country-specific effects (μ_i)

$$\Delta y_{it} = \kappa \Delta y_{i,t-1} + \delta' \Delta W_{it} + \Delta\omega_t + \Delta v_{it} \qquad \text{(II.78)}$$

Assuming the error term to have finite moments and to be serially uncorrelated

$$E(v_{it}) = E(v_{it} v_{is}) = 0 \text{ for } t \neq s \qquad \text{(II.79)}$$

and that the initial conditions satisfy

$$E(y_{i1} v_{it}) = 0 \text{ for } i = 1, \dots, N \text{ and } t = 2, \dots, T \qquad \text{(II.80)}$$

Under these assumptions, second and earlier lagged values of y_{it} are valid instruments for $\Delta y_{i,t-1}$ because the second and earlier lagged values of y_{it} are correlated with $\Delta y_{i,t-1}$ but are not correlated with Δv_{it} . Therefore, the DGMM estimator uses the second and earlier lags of y_{it} as instruments for the equations in first-differences. With respect to X_{it}, depending on X_{it} being treated as exogenous, predetermined or endogenous, the DGMM approach uses suitable lagged values of X_{it} as valid instruments for equations in first differences (Arellano and Bond, 1991; BHT; Hoeffler, 2002; Baltagi, 2008).

If X_{it} are predetermined: $E[X_{it} v_{is}] \neq 0$ for $s < t$ and $E[X_{it} v_{is}] = 0$ for $s \geq t$, and then the first and earlier lagged values of X_{it} are valid instruments for equations in first differences. Namely, for $t = 3, 4, \dots, T$, the model implies the following $(T-2)[(k-1)(T+1) + (T-1)]/2$ moment restrictions

$$E\left[X_{i,t-s}\Delta v_{it}\right] = 0 \text{ for } s = 1, \dots, (t-1) \text{ and } t = 3, \dots, T \qquad (\text{II}.81)$$

The moment restrictions in (II.81) can be written in vector form as follows

$$E[Z_i'\Delta v_i] = 0 \text{ for } i = 1, 2, \dots, N$$

where $\Delta v_i = (\Delta v_{i3}, \Delta v_{i4}, \dots, \Delta v_{iT})'$, Z_i is a $(T-2) \times (T-2)[(k-1)(T+1)+(T-1)]/2$ block diagonal matrix whose s*th* block is given by $\left(y_{i1} \dots y_{is}\, X_{i1}' \dots X_{is}'\, X_{i,s+1}'\right)$ for $s = 1, \dots, T-2$.

If X_{it} are strictly exogenous: $E[X_{it}v_{is}] = 0 \text{ for all } s \text{ and all } t$, then values of X_{it} are themselves valid instruments for equations in first differences. The model implies

$$E[Z_i'\Delta v_i] = 0 \text{ for } i = 1, 2, \dots, N$$

where Z_i is a $(T-2) \times (T-2)[(k-1)(T+1)+(T-1)]/2$ block diagonal matrix whose s*th* block is given by $\left(y_{i1} \dots y_{is}\, X_{i1}' \dots X_{is}'\, X_{i,s+1}'\, X_{i,s+2}'\right)$ for $s = 1, \dots, T-2$.

If X_{it} are endogenous: $E[X_{it}v_{is}] \neq 0 \text{ for } s \leq t$ and $E[X_{it}v_{is}] = 0 \text{ for } s > t$, then the second and earlier lagged values of X_{it} are employed as valid instruments for equations in first differences. The model implies

$$E\ [Z_i'\Delta v_i] = 0 \text{ for } i = 1, 2, \dots, N$$

where Z_i is a $(T-2) \times (T-2)[(k-1)(T+1)+(T-1)]/2$ block diagonal matrix whose s*th* block is given by $(y_{i1} \dots y_{is}\, X_{i1}' \dots\ X_{is}')$ for $s = 1, \dots, T-2$.

Note that X_{it} may include predetermined, exogenous and endogenous variables. In all cases, the first-differenced GMM estimator of the $k \times 1$ coefficient vector δ has the following form

$$\hat{\delta} = (\Delta W' Z A_N Z' \Delta W)^{-1}\, \Delta W' Z A_N Z' \Delta y \qquad (\text{II}.82)$$

where ΔW is a stacked $(T-2)N \times k$ matrix observations on ΔW_{it}, $Z = (Z_1', \dots, Z_N')'$ and $\Delta y = \left(\Delta y_{i2}, \Delta y_{i3}, \dots, \Delta y_{i,T-1}\right)'$.

The choice of the weight matrix A_N in (II.82) yields one-step or two-step estimators as follows.

The two-step estimator $\hat{\delta}_2$ is achieved by using the optimal weight matrix

$$A_{2N} = \left[\frac{1}{N}\sum_{i=1}^{N}\left(Z_i'\widehat{\Delta v_\iota}\widehat{\Delta v_\iota}'Z_i\right)\right]^{-1}$$

where $\widehat{\Delta v_\iota}$ are consistent estimates of residuals achieved from an initial consistent estimator.

Assuming that v_{it} are independent and identically distributed, an asymptotically equivalent GMM estimator $\hat{\delta}_1$ can be achieved in one step using the following weight matrix

$$A_{1N} = \left[\frac{1}{N}\sum_{i=1}^{N}(Z_i'HZ_i)\right]^{-1}$$

where H is a $(T-2) \times (T-2)$ matrix

$$H = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 2 \end{bmatrix}$$

When v_{it} is independently and identically distributed, one-step and two-step estimators are asymptotically equivalent (Blundell and Bond, 1998; Baltagi, 2008).

However, Blundell and Bond (1998) point out that, for finite samples, the lagged levels of highly persistent variables are only weakly correlated with their subsequent first-differences. Therefore, the instruments available for the equations in first differences are weak. This may lead to imprecision and finite sample bias in the estimates of the standard first-differenced GMM estimator. They suggest using a system GMM (SGMM) estimator to deal with the issue of weak instruments.

The SGMM estimator imposes the additional assumption

$$E[\mu_i \Delta y_{i2}] = 0 \text{ for } i = 1, \dots, N \tag{II.83}$$

and

$$E[\mu_i \Delta X_{it}] = 0 \text{ for } i = 1, \dots, N \text{ and } t = 2, \dots, T \tag{II.84}$$

and then the lagged first-differences of y_{it} and X_{it} can be used as valid instruments for the equations in levels (Arellano and Bover, 1995; Blundell and Bond, 1998). The system GMM estimator uses suitable lagged levels of the variables as instruments for the equations in first differences, as in Arellano and Bond, but also adds suitable lagged first-differences of variables as instruments for the equations in levels.

Notice that the assumption that the disturbance v_{it} is not serially correlated can be checked by performing an Arellano-Bond test to test for no second-order serial correlation in the first-differenced residuals. In addition, the validity of instruments used in the first-differenced equations can be checked by using a Hansen test of the overidentifying restrictions while the validity of additional instruments used in levels equations can be tested using a Difference Hansen test (Arellano and Bond, 1991; Blundell and Bond, 1998; Roodman, 2006).

Based on Monte Carlo simulations and asymptotic variance calculations, Blundell and Bond (1998) provided empirical evidence that system GMM can produce dramatic efficiency gains in situations where first-differenced GMM performs poorly. In addition, it is noteworthy that the one-step and two-step first-differenced GMM estimators are equivalent to each other when the error terms are independently and identically distributed across countries and over time. Otherwise the two-step GMM is more efficient than the one-step one and this is always true for system GMM (Arellano and Bond, 1991; Blundell and Bond, 1998; BHT). However, the asymptotic standard errors in the two-step GMM estimators can be downward biased in finite samples (Blundell and Bond, 1998), and thus BHT and Hoeffler (2002) use one-step system GMM to estimate the HCASM. Windmeijer (2005) finds that the downward bias of the asymptotic standard errors in the efficient two-step GMM estimator is due to the estimation of the weight matrix and thus he suggests a correction term which relies on a Taylor series expansion that accounts for the estimation of the weight matrix (Baltagi, 2008). This procedure is advisable because the two-step GMM estimator is more efficient in comparison with the one-step one (Windmeijer, 2005; Roodman, 2006). Therefore, the two-step system GMM with Windmeijer-corrected standard errors (hereafter robust two-step SGMM) is used as a preferred estimation method in this study.

Furthermore, earlier empirical works on the HCASM using the system GMM estimator such as those by BHT and Hoeffler (2002) do not restrict the instrument set. A recent study by Roodman (2009) shows that the use of too many instruments (for endogenous/predetermined variables) in finite samples can overfit endogenous/predetermined variables, bias estimates of the optimal weighting matrix and weaken Hansen and Difference Hansen tests. Therefore, we should test for robustness to reductions in the instrument count (for endogenous/predetermined variables). Roodman suggests using two approaches to reduce the number of instruments. The first approach is to ‘collapse’ the standard instrument matrix by squeezing the matrix horizontally and adding together formerly distinct columns to decrease the number of instruments (Roodman, 2009, p.148). The second approach is to use only certain lags instead of all available lags for instruments. The robust two-step GMM estimation of the HCASM in this study will use these approaches to restrict the instrument count and check robustness.

In the GMM estimators, this study treats initial output per worker and the level of human capital that are measured at the beginning of each five-year period as predetermined variables. The investment rate in physical capital and the population growth rate are treated as endogenous variables because they could be determined by growth and vice versa. For example, it is likely to be a fairly robust stylized fact that growth increases the saving rate (Carroll and Weil, 1994). With imperfect capital mobility, this leads to a relationship between growth and investment even in the absence of any independent causal role for investment (Temple, 1999, p.129). Also, CEL (p.6) assert that there is by now both theoretical and empirical support for the view that economic growth impacts the population growth rate of an economy. In addition, a problem of endogeneity may arise if some omitted variable jointly affects both growth and the explanatory variables (Temple, 1999; Durlauf, Johnson and Temple, 2005).

II.3.2. Testing the significance of sub-Saharan African and East Asian dummies

In order to check whether the HCASM explains the abnormal growth of the output per worker in sub-Saharan Africa and East Asia, this study, in turn, adds dummy variables for sub-Saharan African (SSA) and East Asian (EA) countries into the HCASM and then uses the robust two-step SGMM to estimate the model; the coefficient on the regional dummy can be identified by its inclusion as a strictly exogenous variable in the levels equations of SGMM estimation. In the regression, SSA and EA dummies will take the value equal to one if the country is located in sub-Saharan Africa and East Asia respectively; otherwise the dummies will take the value zero. If the coefficients on the SSA or EA dummies are found to be significant, we can say that the HCASM cannot fully explain the differential in growth in output per worker of the sub-Saharan African or East Asian regions respectively.

Alternatively, we can use a two-step procedure as suggested in Hoeffler (2002) to check whether the HCASM explains the abnormal growth performance of countries in sub-Saharan Africa and East Asia. The procedure includes two steps: the first step estimates the HCASM to compute the residuals, and then the second step will regress the estimated residuals on an SSA or EA dummy. If the coefficient on the SSA or EA dummy in the second step is significant, we can conclude that the HCASM does not fully explain the variation in the growth in sub-Saharan Africa or East Asian regions respectively, compared to other regions. However, different from Hoeffler (2002), this study uses efficient two-step SGMM with Windmeijer-corrected standard errors, instead of the one-step SGMM, to estimate the HCASM in the first step of the procedure.

II.3.3. Sample and data

The data used in this study are collected from the newest Penn World Table version 8.0 recently published by Feenstra, Inklaar and Timmer (2013) (hereafter PWT80) and the newest Educational Attainment Dataset version 1.3 recently published by Barro and Lee (2013) (hereafter BL13). The PWT80 data set consists of 167 countries for 1950-2011 and the BL13 data set consists of 146 countries for 1950-2010. There are 33 countries which are present in the PWT80 but absent in the BL13, whereas there are also 12 countries which are present in the BL13 but absent in the BL13. Hence, when merging the PWT80 data set and the BL13 data set, there are 134 countries where data are available in both data sets. This leads to a sample which comprises 134 countries for 1950-2010 (hereafter Sample1). In order to avoid the potential sample-selection bias induced by data availability, this large Sample 1 is used in this study. In addition, this study sometimes excludes countries for which oil is the dominant industry, making the second sample which comprises 127 countries in the period 1950-2005 (hereafter Sample 2). These oil-producing countries are Bahrain, Gabon, Iran, Iraq, Kuwait, Qatar and Saudi Arabia. Sample 2 is also used in this study. Several empirical papers using OLS estimators often exclude small countries from the sample because the output of the countries may be affected by idiosyncratic factors (MRW, p. 413). However, this study controls for idiosyncratic factors as country-specific effects in the estimation, and thus we do not exclude small countries from our samples. It is noted that data are not available for all countries over the whole period 1950-2005, and thus Sample 1 and 2 used in this study will give rise to unbalanced panels, as in some earlier papers.

In empirical studies, the dependent variable, y_{it}, in the regression human-capital augmented Solow model (II.76) is often measured by the natural logarithm of GDP per capita or the natural logarithm of GDP per worker. For example, MRW use GDP per worker while Islam (1995) uses GDP per capita. However, Durlauf, Jonson and Temple (2005) argue that the (human-capital augmented) Solow model is derived from a production function, and thus it is likely to be appropriate to use GDP per worker rather than GDP per capita because workers are the main force contributing to production. Therefore, this study uses the natural logarithm of GDP per worker. The size of the labour force is proxied by the number of persons engaged (in millions) which is collected from the PWT80. Data on GDP (in millions constant 2005 US$) are also collected from the PWT80.

In this study, the average growth rate of the labour force, n, over each quinquennial period is constructed by dividing the natural log difference in workforce between the beginning and the end of each five-year interval by five. This study follows MRW, Islam (1995) and CEL to choose 0.05 as an assessment of the sum of a common exogenous rate of technological change, g, and a common depreciate rate, δ. However, this study also tried the alternative measure 0.08. There was no considerable difference in the results.

In this study, the proportion of output invested in physical capital, s_k, is proxied by the share of gross capital formation in GDP, averaged over each five-year interval. Data on the gross capital formation share of GDP are collected from the PWT80.

MRW and Islam (1995) show that there are two ways to consider the role of human capital in the growth process, namely by using the rate of investment in human capital or the level of human capital. In the empirical literature, the investment rate in human capital and the level of human capital are often used to investigate the effect of human capital on growth [e.g. MRW and CEL employed the investment rate in human capital whereas Islam (1995) and Hoeffler (2002) employed the level of human capital]. Following Islam (1995) and Hoeffler (2002), this paper uses the level of human capital, h, that is proxied by the average years of total schooling for the population over age 15 at the beginning of each five-year interval. These data are collected from BL13.

In summary, this chapter has discussed the estimator and the data that will be used to test the HCASM. The method used to check whether the HCASM can account for the unusual growth performance for sub-Saharan Africa and East Asia is also presented in the second section. Finally, the samples and data sources used in this study are described in the third section of this chapter. The next chapter will present the empirical results of this study.

CHAPTER II.4:

EMPIRICAL RESULTS

This chapter consists of two sections. The first section presents empirical results of the estimation of the HCASM. The second section reports empirical results of the investigation into whether the HCASM can account for the growth differential of sub-Saharan African and East Asian countries.

II.4.1. Estimating the human-capital augmented Solow model (HCASM)

This section uses the regression model as described in (II.76) in chapter II.3 to test the HCASM. The dependent variable is the natural logarithm of real GDP per worker, $\ln(Y/L)_{it}$. It is regressed on the natural logarithm of initial real GDP per worker, $\ln(Y/L)_{i,t-1}$, the natural logarithm of the investment rate, $\ln s_{it}$, the natural logarithm of the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation), $\ln(n_{it} + g + \delta)$, and the average years of total schooling measured at the beginning of each five-year period, $\ln h_{it}$. The econometric package used is STATA. All the regressions include period dummies.

Sample 1 is firstly used to test the HCASM. Table II.4.1 reports the results of regressions using Pooled OLS (POLS), Within Groups (WG), two-step first-differenced GMM estimation with Windmeijer-corrected standard errors (hereafter DGMM2R) and an efficient two-step system GMM estimation with Windmeijer-corrected standard errors (hereafter SGMM2R).

The DGMM2R in Table II.4.1 uses a standard instrument set suggested for conventional DGMM by Arellano and Bond (1991) (hereafter DGMM2R_SIS). In particular, it uses $\ln(Y/L)_{i,t-2}$, $\ln s_{i,t-2}$, $\ln\big(n_{i,t-2} + g +$

$\delta)$ and $\ln h_{i,t-1}$ and all earlier lags as instruments for first differenced equations.

Table II.4.1: Estimation of the HCASM – Sample 1

	(1) POLS	(2) WG	(3) DGMM2R_SIS	(4) SGMM2R_SIS
$\ln(Y/L)_{i,t-1}$	0.961*** (0.008)	0.770*** (0.035)	0.732*** (0.059)	0.943*** (0.013)
$\ln s_{it}$	0.073*** (0.016)	0.088*** (0.025)	0.069** (0.029)	0.106*** (0.025)
$\ln(n_{it}+g+\delta)$	-0.105*** (0.024)	-0.109*** (0.030)	-0.097*** (0.036)	-0.149*** (0.041)
$\ln h_{it}$	0.043*** (0.017)	-0.093*** (0.036)	-0.123** (0.059)	0.058** (0.024)
Implied λ F-test (p-value)	0.008 (0.002) 0.30	0.052 (0.009) 0.61	0.062 (0.016) 0.54	0.012 (0.003) 0.79
m2 (p-value)	-	-	0.37	0.44
Hansen (p-value)	-	-	1.00	1.00
Dif.Hansen (p-value)	-	-	-	1.00
No. of instr.$^{(a)}$	-	-	242	285
No. of observations	1170	1170	1032	1170
No. of countries	134	134	134	134
Average no. of periods	8.7	8.7	7.7	8.7

Note: *** denotes significance at the one percent level. ** denotes significance at the five percent level. * denotes significance at the ten percent level. Standard errors are reported in parentheses. $(Y/L)_{i,t-1}$ denotes initial output per worker, s_{it} investment rate, $(n_{it}+g+\delta)$ growth rate of labour (adjusted by the rate of technological progress and the rate of depreciation) and h_{it} the level of human capital. λ denotes convergence rate. (a) the number of instruments used for endogenous and predetermined variables. POLS and FE use robust standard errors. DGMM2R_SIS and SGMM2R_SIS use Windmeijer-corrected standard errors. F-test is used to test the restriction that the coefficients on the physical investment rate and the growth of labour (adjusted by the rate of technological progress and the rate of depreciation) have the same magnitude and opposite signs. m_2 is Arellano-Bond test of second order serial correlation. STATA output of the regressions is reported in appendix II.5.

The SGMM2R in Table II.4.1 uses a standard instrument set suggested for conventional SGMM by Blundell and Bond (1998) (hereafter SGMM2R_SIS). Namely, it uses $\ln(Y/L)_{i,t-2}$, $\ln s_{i,t-2}$, $\ln(n_{i,t-2} + g + \delta)$ and $\ln h_{i,t-1}$ and all earlier lags as instruments for first differenced equations, combined with $\Delta \ln(Y/L)_{i,t-1}$, $\Delta \ln s_{i,t-1}$, $\Delta \ln(n_{i,t-1} + g + \delta)$ and $\Delta \ln h_{it}$ as instruments for level equations.

Firstly, in the POLS regression, which uses a conventional least squares regression based on pooling all observations without considering country-specific effects, the coefficient on the variable for initial output per worker, $\ln(Y/L)_{i,t-1}$, is found to be significant at the one percent level with a value less than one. This implies an existence of conditional convergence as being predicted in the HCASM. The estimated rate of convergence, which is derived from the coefficient on initial output per worker, is approximately 0.8 percent per year (λ = 0.8%). The OLS regression in the first column of Table II.4.1 also shows that the coefficients on the variables for the investment rate in physical capital, $\ln s_{it}$, and the level of human capital, $\ln h_{it}$, are found to be significant and positive while that on the growth of the labour force (adjusted by the rate of technological progress and the rate of depreciation), $\ln(n_{it} + g + \delta)$, is found to be significant and negative. These results support the HCASM in that the investment rate and the level of human capital impact output per worker positively while labour force growth impacts output per worker negatively. Further, the result of an F-test does not reject the restriction in the HCASM that the coefficients on the two variables for the investment rate in physical capital and the growth of the labour force (adjusted by the rate of technological progress and the rate of depreciation) are equal in magnitude but opposite in sign.

In the WG regression in the second column of Table II.4.1, when country-specific effects are controlled for, the coefficient on the variable for initial output per worker, $\ln(Y/L)_{i,t-1}$, is also found to be less than one and significant, indicating an existence of conditional convergence as expected in the HCASM. However, the magnitude of the coefficient is found to be smaller than that in the POLS results, and thus the convergence rate found from the WG regression is higher than that estimated from the POLS. It is around 5.2 percent a year ($\lambda = 5.2\ \%$). Variables for the investment rate in physical capital, $\ln s_{it}$, and the growth of the labour force (adjusted by the rate of technological progress and the rate of depreciation), $\ln(n_{it} + g + \delta)$, are found to have significant positive and negative effects, respectively, as predicted in the HCASM. In addition, the restriction in the HCASM that the coefficients of these two variables have the same magnitude and opposite signs is not rejected by the F-test. However, the coefficient on the variable for human capital, $\ln h_{it}$, is found to be significant with the wrong sign. The result is similar to that in the research by Islam (1995) and CEL where they also uses fixed-effects estimations and finds a negative and significant coefficient on human capital. Note that the WG estimator is known to be likely to be biased in the presence of the lagged dependent variable (Nickell, 1981), especially in a sample with large N and small T as in this research. In addition, it does not address the potential endogeneity of the regressors. Therefore, these limitations may cause the significance of the human capital with the wrong sign.

Furthermore, as discussed in chapter II.3, the estimate of the coefficient on the lagged dependent variable, $\ln(Y/L)_{i,t-1}$, in the POLS is likely to be biased upwards while that in the WG is likely to be biased downwards. Therefore, the two estimates here may be useful since they provide the approximate upper and lower bounds for other estimations to some extent, as in BHT. One might hope that a consistent estimate (of the coefficient on the lagged dependent variable) should lie between the OLS and WG estimates or should not be significantly higher than the former or significantly lower than the latter (Bond, 2002). Note that the bounds are only approximate because of the sample variability and presence of explanatory variables other than the lagged dependent variable, such as the investment rate, the growth rate of the labour force and the level of human capital. In addition, the explanatory variables are instrumented in the GMM estimations, but not in the POLS and WG estimations.

With respect to the DGMM2R and SGMM2R estimations using a standard instrument sets in column three of Table II.4.1, it can be seen that the number of instruments (for endogenous and predetermined variables) in these GMM estimations is large. The DGMM2R_SIS regression uses 242 instruments while the SGMM2R_SIS regression uses 285 instruments. These numbers of instruments used are much larger than the number of countries in the sample (*N*=134). Roodman (2009) shows that the use of too many instruments (for endogenous/predetermined variables) may overfit endogenous/predetermined variables, bias estimates of the optimal weighting matrix and weaken Hansen and Difference Hansen tests. Clearly, the Hansen and Hansen tests the DGMM2R_SIS and SGMM2R_SIS regression give implausibly perfect p-values. Therefore, these tests are likely to be weakened and thus we do not know whether the instrument sets used in the DGMM2R_SIS and SGMM2R_SIS regressions from Table II.4.1 are potentially valid or not. Therefore, this study experiments to reduce the number of instrument by ‘collapsing’ the standard instrument set or using one or several lags instead of all available lags to choose valid instrument sets for DGMM2R and SGMM2R estimation.

Table II.4.2: Experiment instrument sets in the DGMM2R estimation – Sample 1

Instrument set	
1. 'Collapse' the standard instrument set (used in the DGMM2R regression in Table II.4.1)	Hansen (p-value): 0.02 No. of instr.$^{(a)}$: 42
2. Use $\ln(Y/L)_{i,t-2}$, $\ln s_{i,t-2}$, $\ln(n_{i,t-2}+g+\delta)$ and $\ln h_{i,t-1}$ and as instruments.	Hansen (p-value): 0.03 No. of instr.$^{(a)}$: 42
3. Use $\ln(Y/L)_{i,t-2}$ and $\ln(Y/L)_{i,t-3}$, $\ln s_{i,t-2}$ and $\ln s_{i,t-3}$, $\ln(n_{i,t-2}+g+\delta)$ and $\ln(n_{i,t-3}+g+\delta)$, and $\ln h_{i,t-1}$ and $\ln h_{i,t-2}$ as instruments.	Hansen (p-value): 0.08 No. of instr.$^{(a)}$: 80
4. Use $\ln(Y/L)_{i,t-2}$ ... $\ln(Y/L)_{i,t-4}$, $\ln s_{i,t-2}$ ... $\ln s_{i,t-4}$, $\ln(n_{i,t-2}+g+\delta)$… $\ln(n_{i,t-4}+g+\delta)$, and $\ln h_{i,t-1}$ … $\ln h_{i,t-3}$ as instruments.	Hansen (p-value): 0.31 No. of instr.$^{(a)}$: 114
5. Use $\ln(Y/L)_{i,t-3}$, $\ln s_{i,t-3}$, $\ln(n_{i,t-3}+g+\delta)$ and $\ln h_{i,t-2}$ and as instruments.	Hansen (p-value): 0.02 No. of instr.$^{(a)}$: 38
6. Use $\ln(Y/L)_{i,t-3}$ and $\ln(Y/L)_{i,t-4}$, $\ln s_{i,t-3}$ and $\ln s_{i,t-4}$, $\ln(n_{i,t-3}+g+\delta)$ and $\ln(n_{i,t-4}+g+\delta)$, and $\ln h_{i,t-2}$ … $\ln h_{i,t-3}$ as instruments.	Hansen (p-value): 0.02 No. of instr.$^{(a)}$: 72
7. Use $\ln(Y/L)_{i,t-4}$, $\ln s_{i,t-4}$, $\ln(n_{i,t-4}+g+\delta)$ and $\ln h_{i,t-3}$ and as instruments.	Hansen (p-value): 0.01 No. of instr.$^{(a)}$: 34
8. Use $\ln(Y/L)_{i,t-4}$ and $\ln(Y/L)_{i,t-5}$, $\ln s_{i,t-4}$ and $\ln s_{i,t-5}$, $\ln(n_{i,t-4}+g+\delta)$ and $\ln(n_{i,t-5}+g+\delta)$, and $\ln h_{i,t-3}$ … $\ln h_{i,t-4}$ as instruments.	Hansen (p-value): 0.07 No. of instr.$^{(a)}$: 64

(a) The number of instruments used for endogenous and predetermined variables.

Instrument sets are experimented with to reduce the instrument count in the DGMM2R estimates, with results that are presented in Table II.4.2. The left column of this table describes the instrument set used in each DGMM2R regression while the right column reports the corresponding result of a Hansen test and the number of instruments. The Hansen test is used to check the validity of instruments.

Roodman (2006, p.44) notes that 'do not take comfort in a Hansen test p value somewhat above 0.1'. Among the instrument sets experimented with in Table II.4.2, only instrument set 4 is not rejected by a Hansen test (p-value=0.31). Other instrument sets in Table II.4.2 are rejected by the results of Hansen tests. The instrument count in the DGMM2R regression using instrument set 4 [i.e. using $\ln(Y/L)_{i,t-2}$ ... $\ln(Y/L)_{i,t-4}$, $\ln s_{i,t-2}$ ... $\ln s_{i,t-4}$, $\ln\left(n_{i,t-2} + g + \delta\right)$... $\ln\left(n_{i,t-4} + g + \delta\right)$, and $\ln h_{i,t-1}$ … $\ln h_{i,t-3}$ as instruments] is 114 which is much smaller than in the DGMM2R using a standard instrument set as in Table II.4.1 (242). It is also smaller than the number of countries in Sample 1 (*N*=134). The results of the DGMM2R regression using instrument set 4 are reported in column one of Table II.4.4.

With respect to the SGMM2 estimation, results of experiments to reduce the instrument count are presented in Table II.4.3. The left column of this table describes the instrument set used in each SGMM2R regression while the right column presents the corresponding result of Hansen and Difference Hansen tests and the number of instruments.

Table II.4.3:
Experiment instrument sets in the robust SGMM2R estimation – Sample 1

Instrument set	
1. 'Collapse' the standard instrument set (used in the SGMM2R regression in Table II.4.1)	Hansen (p-value): 0.01 Dif. Hansen (p-value): 0.04 No. of instr.$^{(a)}$: 47
2. Use $\ln(Y/L)_{i,t-2}$, $\ln s_{i,t-2}$, $\ln(n_{i,t-2}+g+\delta)$ and $\ln h_{i,t-1}$ and as instruments for first differenced equations; combined with $\Delta\ln(Y/L)_{i,t-1}$, $\Delta\ln s_{i,t-1}$, $\Delta\ln(n_{i,t-1}+g+\delta)$ and $\Delta\ln h_{it}$ as instruments for level equations.	Hansen (p-value): 0.03 Dif. Hansen (p-value): 0.14 No. of instr.$^{(a)}$: 85
3. Use $\ln(Y/L)_{i,t-2}$ and $\ln(Y/L)_{i,t-3}$, $\ln s_{i,t-2}$ and $\ln s_{i,t-3}$, $\ln(n_{i,t-2}+g+\delta)$ and $\ln(n_{i,t-3}+g+\delta)$, and $\ln h_{i,t-1}$ and $\ln h_{i,t-2}$ as instruments for first differenced equations; combined with $\Delta\ln(Y/L)_{i,t-1}$, $\Delta\ln s_{i,t-1}$, $\Delta\ln(n_{i,t-1}+g+\delta)$ and $\Delta\ln h_{it}$ as instruments for level equations.	Hansen (p-value): 0.27 Dif. Hansen (p-value): 0.84 No. of instr.$^{(a)}$: 123
4. Use $\ln(Y/L)_{i,t-2}$ ... $\ln(Y/L)_{i,t-4}$, $\ln s_{i,t-2}$ ... $\ln s_{i,t-4}$, $\ln(n_{i,t-2}+g+\delta)\ldots\ \ln(n_{i,t-4}+g+\delta)$, and $\ln h_{i,t-1}\ \ldots\ \ln h_{i,t-3}$ as instruments for first differenced equations; combined with $\Delta\ln(Y/L)_{i,t-1}$, $\Delta\ln s_{i,t-1}$, $\Delta\ln(n_{i,t-1}+g+\delta)$ and $\Delta\ln h_{it}$ as instruments for level equations.	Hansen (p-value): 0.95 Dif. Hansen (p-value): 1.00 No. of instr.$^{(a)}$: 157
5. Use $\ln(Y/L)_{i,t-3}$, $\ln s_{i,t-3}$, $\ln(n_{i,t-3}+g+\delta)$ and $\ln h_{i,t-2}$ and as instruments for first differenced equations; combined with $\Delta\ln(Y/L)_{i,t-2}$, $\Delta\ln s_{i,t-2}$, $\Delta\ln(n_{i,t-2}+g+\delta)$ and $\Delta\ln h_{i,t-1}$ as instruments for level equations.	Hansen (p-value): 0.14 Dif. Hansen (p-value): 0.84 No. of instr.$^{(a)}$: 77
6. Use $\ln(Y/L)_{i,t-3}$ and $\ln(Y/L)_{i,t-4}$, $\ln s_{i,t-3}$ and $\ln s_{i,t-4}$, $\ln(n_{i,t-3}+g+\delta)$ and $\ln(n_{i,t-4}+g+\delta)$, and $\ln h_{i,t-2}$ and $\ln h_{i,t-3}$ as instruments for first differenced equations; combined with $\Delta\ln(Y/L)_{i,t-2}$, $\Delta\ln s_{i,t-2}$, $\Delta\ln(n_{i,t-2}+g+\delta)$ and $\Delta\ln h_{i,t-1}$ as instruments for level equations.	Hansen (p-value): 0.27 Dif. Hansen (p-value): 0.48 No. of instr.$^{(a)}$: 111
7. Use $\ln(Y/L)_{i,t-4}$, $\ln s_{i,t-4}$, $\ln(n_{i,t-4}+g+\delta)$ and $\ln h_{i,t-3}$ and as instruments for first differenced equations; combined with $\Delta\ln(Y/L)_{i,t-3}$, $\Delta\ln s_{i,t-3}$, $\Delta\ln(n_{i,t-3}+g+\delta)$ and $\Delta\ln h_{i,t-2}$ as instruments for level equations.	Hansen (p-value): 0.01 Dif. Hansen (p-value): 0.12 No. of instr.$^{(a)}$: 69
8. Use $\ln(Y/L)_{i,t-4}$ and $\ln(Y/L)_{i,t-5}$, $\ln s_{i,t-4}$ and $\ln s_{i,t-5}$, $\ln(n_{i,t-4}+g+\delta)$ and $\ln(n_{i,t-5}+g+\delta)$, and $\ln h_{i,t-3}\ \ldots\ \ln h_{i,t-4}$ as instruments for first differenced equations; combined with $\Delta\ln(Y/L)_{i,t-3}$, $\Delta\ln s_{i,t-3}$, $\Delta\ln(n_{i,t-3}+g+\delta)$ and $\Delta\ln h_{i,t-1}$ as instruments for level equations.	Hansen (p-value): 0.13 Dif. Hansen (p-value): 0.24 No. of instr.$^{(a)}$: 99

(a) The number of instruments used for endogenous and predetermined variables.

Among instrument sets experimented with as summarised in Table II.4.3, only instruments sets 3, 5 and 6 are not rejected by Hansen and Difference Hansen test at the conventional levels. However, the p-value in the Hansen test for instrument set 5 is just 0.14, and thus the validity of this instrument set seems not to be safely confirmed. As for instrument sets 3 and 6, Hansen and Difference Hansen tests imply that the instruments included in these sets are potentially valid (there is no evidence to reject their validity). The instrument count in the SGMM2 regression using instrument set 3 is 123 while that in the SGMM2R using set 6 is 111. These instrument counts are much smaller than in the SGMM2R regression using a standard instrument set in column four of Table II.4.1 (285), and they are also smaller than the number of countries of the sample (N=134). Results of the SGMM2 estimations using instrument set 3 and 6 are reported in column two and three of Table II.4.4.

The results of the DGMM2R estimation using $\ln(Y/L)_{i,t-2}$... $\ln(Y/L)_{i,t-4}$, $\ln s_{i,t-2}$ ... $\ln s_{i,t-4}$, $\ln(n_{i,t-2} + g + \delta)$... $\ln(n_{i,t-4} + g + \delta)$, and $\ln h_{i,t-1}$ ... $\ln h_{i,t-3}$ as valid instruments in column one of Table II.4.4 show that the estimated coefficient on the variable of initial output per worker, $\ln(Y/L)_{i,t-1}$, is significant and equal to 0.61, indicating conditional convergence as predicted in the HCASM. The estimated rate of convergence is found to be high, approximately 9.8 per cent a year. The investment rate is found to be weakly significant with the expected sign while the human capital level is found to be significant with the wrong (negative) sign. However, it can be seen that the estimate of the coefficient on $\ln(Y/L)_{i,t-1}$ in the DGMM2R regression seems to be much lower than that in the corresponding WG estimate (see column two of Table II.4.1) which itself is likely to be biased downwards. Therefore, the DGMM2R estimate of the coefficient on initial output per worker is likely to be biased. This is consistent with the results of simulations in econometric studies, for example Blundell and Bond (1998), which indicate that the first-differenced GMM estimator may be seriously biased due to the presence of weak instruments. Therefore, the finding on the significance of the human capital with a wrong sign may be a consequence of the presence of bias in first-differenced GMM regressions, and overall these results should be treated with caution.

Next, this study considers the result of the SGMM2R with valid instrument sets 3 and 6 in columns two and three of Table II.4.4. In particular, column two reports SGMM2R estimates using $\ln(Y/L)_{i,t-2}$ and $\ln(Y/L)_{i,t-3}$, $\ln s_{i,t-2}$ and $\ln s_{i,t-3}$, $\ln\left(n_{i,t-2} + g + \delta\right)$ and $\ln\left(n_{i,t-3} + g + \delta\right)$, and $\ln h_{i,t-1}$ and $\ln h_{i,t-2}$ as valid instruments for first differenced equations, combined with $\Delta \ln(Y/L)_{i,t-1}$, $\Delta \ln s_{i,t-1}$, $\Delta \ln\left(n_{i,t-1} + g + \delta\right)$ and $\Delta \ln h_{it}$ as valid instruments for level equations (i.e. instrument set 3); whereas column three report SGMM2R estimates using $\ln(Y/L)_{i,t-3}$ and $\ln(Y/L)_{i,t-4}$, $\ln s_{i,t-3}$ and $\ln s_{i,t-4}$, $\ln\left(n_{i,t-3} + g + \delta\right)$ and $\ln\left(n_{i,t-4} + g + \delta\right)$, and $\ln h_{i,t-2}$ and $\ln h_{i,t-3}$ as valid instruments for first differenced equations combined with $\Delta \ln(Y/L)_{i,t-2}$, $\Delta \ln s_{i,t-2}$, $\Delta \ln\left(n_{i,t-2} + g + \delta\right)$ and $\Delta \ln h_{i,t-1}$ as valid instruments for level equations (i.e. instrument set 6).

Table II.4.4: The DGMM2R, SGMM2R and LSDVC estimation of the HCASM – Sample 1

	(1) DGMM2R_IS4	(2) SGMM2R_IS3	(3) SGMM2R_IS6	(4) LSDVC1	(5) LSDVC2
$\ln(Y/L)_{i,t-1}$	0.613*** (0.078)	0.949*** (0.017)	0.969*** (0.015)	0.997*** (0.021)	0.987*** (0.021)
$\ln s_{it}$	0.051* (0.030)	0.111*** (0.030)	0.088*** (0.028)	0.082** (0.034)	0.082*** (0.029)
$\ln(n_{it} + g + \delta)$	-0.040 (0.063)	-0.207*** (0.062)	-0.134* (0.071)	-0.119*** (0.038)	-0.113*** (0.033)
$\ln h_{it}$	-0.212*** (0.080)	0.033 (0.029)	0.030 (0.035)	-0.033 (0.069)	-0.039 (0.059)
Implied λ	0.098 (0.026)	0.010 (0.004)	0.006 (0.003)	0.001 (0.004)	0.003 (0.004)
F-test (p-value)	0.866	0.15	0.47	0.45	0.47
m2 (p-value)	0.31	0.44	0.44	-	-
Hansen (p-value)	0.31	0.27	0.27	-	-
Dif.Hansen (p-value)	-	0.84	0.48	-	-
No. of instr.$^{(a)}$	114	123	111	-	-
No. of observations	1032	1170	1170	1170	1170
No. of countries	134	134	134	134	134
Average no. of periods	7.7	8.7	8.7	8.7	8.7

Note: *** denotes significance at the one percent level. ** denotes significance at the five percent level. * denotes significance at the ten percent level. Windmeijer-corrected standard errors are reported in parentheses. $(Y/L)_{i,t-1}$ denotes initial output per worker, s_{it} physical investment rate, $(n_{it} + g + \delta)$ growth rate of labour (adjusted by the rate of technological progress and the rate of depreciation) and h_{it} the level of human capital. λ denotes convergence rate. (a) the number of instruments used for endogenous and predetermined variables. F-test is used to test the restriction that the coefficients on the physical investment rate and the growth of labour (adjusted by the rate of technological progress and the rate of depreciation) have the same magnitude and opposite signs. m_2 is Arellano-Bond test of second order serial correlation. DGMM2R_IS4 denotes DGMM2R using instrument set 4 in Table II.4.2. SGMM2R_IS3 and SGMM2R_IS6 denote SGMM2R using instrument sets 3 and 6 in Table II.4.3 respectively. LSDVC1 and LSDVC1 denote LSDVC initialized with SGMM2R_IS3 and SGMM2R_IS6 respectively. STATA output of the regressions is reported in appendix II.6.

Results of the SGMM2R estimations support the prediction of conditional convergence in the HCASM. The coefficients on the initial output per worker, $\ln(Y/L)_{i,t-1}$, in the two SGMM2R estimations using instrument sets 3 and 6 are found to be significant and equal to 0.95 and 0.97 respectively, leading to slow rates of conditional convergence, as in BHT (2001) but here using a reduced instrument set, more countries and a longer span of data. The estimate of the coefficient in the SGMM2R regression using instrument set 3 (hereafter SGMM2R_IS3) lies within the approximate 'upper and lower bounds', but implies a relatively high coefficient on the variable $\ln(Y/L)_{i,t-1}$, similar to that found by the SGMM2R regression using instrument set 6 (hereafter SGMM2R_IS6). Although the estimate of the coefficient in the SGMM2R regression using instrument set 6 lies outside the approximate bounds, it is very close to the (approximate) upper bound. Note that the 'upper and lower bounds' are only approximate as discussed above. Arellano-Bond tests do not reject the null hypothesis of no second-order serial correlation in the regression models in the two estimations. Results of Hansen and Difference Hansen tests imply that the instrument sets used in the two SGMM2R estimations in Table II.4.4 are potentially valid (and the instrument counts are smaller than the number of countries, *N,* of the sample). Therefore, the SGMM2R_IS3 and SGMM2R_IS6 estimates are used as the preferred estimators for Sample 1 in this study.

The rate of convergence rate found in the SGMM2R_IS3 estimation is about one per cent a year ($\lambda = 1\%$) while that found in the SGMM2R_IS6 is approximately 0.6 per cent a year ($\lambda = 0.6\%$). With respect to explanatory variables, the coefficients on the investment rate in physical capital, $\ln s_{it}$, are strongly significant and positive in both the two SGMM2R estimations. This result supports the HCASM in that the effect of the investment rate in physical capital on output per worker is positive. The effect of the growth of the labour force on output per worker is found to be negative as the HCASM predicts. This effect is significant at the one percent level of significance in the SGMM2R_IS3 estimates while it is significant at the ten percent level in the SGMM2R_IS6 estimates. F-tests do not reject the restriction that the coefficients on the investment rate in physical capital and the growth of the labour force have the same magnitude and opposite signs at conventional levels, consistent with the HCASM.

Differentiating from the SGMM2R estimation using a standard instrument set in Table II.4.1, both the SGMM2R_IS3 and SGMM2R_IS6 estimations in Table II.4.4 show that the effect of the human capital on output is insignificant, but positively signed and with p-values of 0.25 and 0.40 respectively. This result is similar to that found in the system GMM estimation of the HCASM in BHT.

Further, this study performs LSDVC estimations initialized with the SGMM2R_IS3 and SGMM2R_IS6 estimates to correct the Nickell (1981) bias in WG estimation. The LSDVC estimations (see column four and five of Table II.4.4) show that the coefficient on the variable for the investment rate, $\ln s_{it}$, is positive and significant while that on the growth of the labour force (adjusted by the rate of technological progress and the rate of depreciation), $\ln(n_{it} + g + \delta)$, is negative and significant at conventional levels. This result supports the prediction of the HCASM that the investment rate in physical capital has a positive effect on output per worker whereas the growth rate of the labour force has a negative effect on output per worker. The coefficient on the level of human capital, $\ln h_{it}$, is found to be insignificant. The result of an F-test does not reject the restriction that the coefficients on the investment rate in physical capital and the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation) have the same magnitude but are opposite in sign, as expected in the HCASM. In addition, the results of the LSDVC estimations support the conditional convergence prediction of the HCASM. The coefficients on the initial output per worker are found to be high, leading to slow rates of conditional convergence. In particular, the rates of convergence, derived from these estimated coefficients on initial output per worker in the LSDVC estimations, are about 0.1 and 0.3 percent a year.

It can be seen that the LSDVC estimates seem in line with our preferred SGMM2R estimates, in that the influence of the investment rate and the growth rate of the labour force (adjusted by the rate of technological progress and the rate of depreciation) on the level of the growth path is significant, while that of the level of human capital is insignificant. They also suggest the rate of conditional convergence is very low. However, note that the LSDVC estimator assumes that the explanatory variables (other than the lagged dependent variable) are strictly exogenous. This assumption seems to be too strong in our case, and thus the SGMM2R estimator is still the preferred one.

This study also tests the robustness of the estimation results of the HCASM by using SGMM2R regressions for Sample 2. Instrument sets in Table II.4.3 are experimented with (see appendix II.4). Among them, only instrument set 6 is likely to be valid based on the results of Hansen and Difference Hansen tests and the comparison between the number of instruments (used for endogenous/predetermined variables) and the number of countries, N, of the sample. Therefore, the robust efficient two-step system GMM regression using instrument 6 (SGMM2R_IS6) is used as the preferred estimator for Sample 2.

Table II.4.5:

The SGMM2R estimation of the HCASM with valid instrument set – Sample 2

	SGMM2R_IS6
$\ln(Y/L)_{i,t-1}$	0.988*** (0.014)
$\ln s_{it}$	0.075*** (0.029)
$\ln(n_{it} + g + \delta)$	0.026 (0.050)
$\ln h_{it}$	0.042 (0.037)
Implied λ F-test (p-value)	0.003 (0.003) 0.05
m2 (p-value)	0.21
Hansen (p-value)	0.38
Dif.Hansen (p-value)	0.60
No. of instr.$^{(a)}$	111
No. of observations	1114
No. of countries	127
Average no. of periods	8.8

Note: *** denotes significance at the one percent level. ** denotes significance at the five percent level. * denotes significance at the ten percent level. Windmeijer-corrected standard errors are reported in parentheses. $(Y/L)_{i,t-1}$ denotes initial output per worker, s_{it} physical investment rate, $(n_{it} + g + \delta)$ growth rate of labour (adjusted by the rate of technological progress and the rate of depreciation) and h_{it} the level of human capital. λ denotes convergence rate. F-test is used to test the restriction that the coefficients on the physical investment rate and the growth of labour (adjusted by the rate of technological progress and the rate of depreciation) have the same magnitude and opposite signs. m_2 is Arellano-Bond test of second order serial correlation. SGMM2R_IS6 denotes SGMM2R using instrument set 6 in Table II.4.3. STATA output of the regression is reported in appendix II.7.

The results of the SGMM2R_IS6 estimates in Table II.4.5 show the coefficient on the variable for initial output per worker is significant and less than one. This indicates conditional convergence as the HCASM predicts. The rate of conditional convergence is approximately 0.3 per cent a year ($\lambda = 0.3\%$). The investment rate is found to be highly significant and positive. The growth of the labour force and the level of human capital are found to be insignificant in the SGMM2R_IS6 regression using Sample 2. The F-test rejects the restriction that the coefficients on the physical investment rate and the growth rate of the labour force have the same magnitude and opposite signs.

In summary, the coefficient on initial output per worker is significantly different from zero, and less than one, in all regressions in this section, robustly confirming the conditional-convergence prediction in the HCASM. The investment rate is also found to be significant and positive in all regressions. This evidence supports the prediction of HCASM that the investment rate has a positive impact on output per worker. In our preferred regressions (SGMM2R_IS3 and SGMM2R_IS6 for Sample 1 and SGMM2R_IS6 for Sample 2), the rate of convergence rate found is from 0.03 to 1 percent a year. These rates of convergence are slower than found in the empirical estimation of the HCASM in previous research by MRW ($\lambda = 1.4\,\%$), Islam (1995) ($\lambda = 3.8\%$), CEL ($\lambda = 6.8\%$), BHT ($\lambda = 1.7\%$) and Hoeffler (2002) ($\lambda = 3.3\%$). The influence of the growth of the labour force on output is not robustly significant. It is found to be significant at the one percent level in the SGMM2R_IS3 using Sample 1 and weakly significant at the ten percent level in the SGMM2R_IS6 using Sample 1. However, it is not significant in the SGMM2R_IS6 using Sample 2. The effect of the human capital variable is found to be insignificant in the preferred SGMM2R regressions. This is a common finding in much of the literature that estimates cross-country panels with education data (Pritchett, 1996; Topel, 1999).

II.4.2. Does the HCASM fully account for the differential growth of output per worker for sub-Saharan African and East Asian countries?

In this section, we investigate whether the HCASM accounts for the unusual growth of countries in sub-Saharan African (SSA) and East Asian (EA) - the two regions are often known as contrasting regions in terms of economic growth over the past fifty years. Descriptive statistics in Table II.4.6 show that eight out of the top ten slowest-growth countries for the period from 1950/1960 to 2010 come from sub-Saharan Africa. The eight countries are Congo D.R., Niger, Zimbabwe, Senegal, Zambia, Ghana, South Africa and Cote d`Ivoire. By contrast, five within the top ten countries with the highest growth rates over the past fifty years are EA countries. They are China, Taiwan, Korea, Singapore and Thailand (see Table II.4.7).

Our calculation based on data from PWT80 finds that the GDP per worker of SSA countries grows, on average, at approximately 0.3 per cent a year while that of EA countries grows, on average, at about 4.7 percent a year, for the period 1960-2010. The average growth rate of GDP per worker of countries in other regions in our sample is found to be around 1.7 per cent a year in the corresponding period. Therefore, it is likely that SSA countries experience a slower growth performance while, conversely, EA countries experience a faster growth performance in comparison with countries in other regions for the past fifty years.

Table II.4.6: Some descriptive statistics of the ten countries with the slowest growth rates for 1950/1960-2010

	Per worker GDP growth (a)	Initial GDP per worker (2005US$)(b)	Average ratio of investment to GDP(a)	Average growth rate of labour (a)	Initial years of schooling (b)
Congo D.R.	-0.0237	2273	0.2543	0.0282	0.81
Niger	-0.0146	3534	0.2290	0.0340	0.38
Zimbabwe	-0.0073	10032	0.1627	0.0319	2.52
Senegal	-0.0009	3857	0.085	0.0275	1.85
Zambia	0.0002	3222	0.1057	0.0267	2.26
Jamaica	0.0019	10004	0.1778	0.0144	3.84
Ghana	0.0029	4063	0.2275	0.0283	1.08
South Africa	0.0041	17010	0.2544	0.0271	4.40
Bolivia	0.0056	6728	0.1307	0.0186	2.50
Cote d`Ivoire	0.0056	2930	0.0889	0.0319	0.93

(a) Statistics are measured from the earliest available years (in 1950-1960) to 2010. (b) Statistics are measured from the earliest available years.

Table II.4.7: Some descriptive statistics of the top ten countries with the fastest growth rates for 1950/1960-2010

	Per worker GDP growth (a)	Initial GDP per worker (2005US$) (b)	Average ratio of investment to GDP (a)	Average growth rate of labour (a)	Initial years of schooling (b)
China	0.0526	817	0.2611	0.0227	1.86
Romania	0.0487	2249	0.2039	-0.0071	5.24
Taiwan	0.0476	6661	0.2625	0.0232	4.98
Korea	0.0461	5842	0.2843	0.0253	4.34
Thailand	0.0460	1478	0.2534	0.0203	4.19
Japan	0.0422	5171	0.2936	0.0084	6.86
Malta	0.0388	7201	0.3639	0.0113	4.17
Singapore	0.0386	10365	0.4567	0.0358	3.66
Turkey	0.0375	4335	0.2170	0.0102	1.11
Cyprus	0.0366	9085	0.3851	0.0107	5.07

(a) statistics are measured from the earliest available years (in 1950-1960) to 2010. (b) statistics are measured from the earliest available years.

This study firstly starts the investigation on whether the HCASM fully explains the growth of output per worker in SSA and EA countries by adding a SSA dummy and an EA dummy to the HCASM and using our preferred robust SGMM2R estimators (SGMM2R_IS3 and SGMM2R_IS6 for Sample 1 and SGMM2R_IS6 for Sample 2) to estimate. In the regressions, the SSA dummy and the EA dummy take the value equal to one if the country is located in sub-Saharan Africa and East Asia respectively; otherwise the dummies take the value zero. The dummies are included in the levels equations of the SGMM estimator.

The SGMM2R results for Samples 1 and 2 are reported in Tables II.4.8 and II.4.9 respectively. In the tables, the estimated models in regressions one and four are the HCASM with an SSA dummy, in regressions two and five are the HCASM with an EA dummy, and in regressions three and six are HCASM with SSA and EA dummies. In Table II.4.8, regressions one, two and three use the SGMM2R_IS3 while regressions four, five and six use the SGMM2R_IS6. In Table II.4.9 regressions one, two and three use the SGMM2R_IS6.

Table II.4.8: SGMM2R estimation of the HCASM with SSA and EA dummies – Sample 1

	SGMM2R_IS3 (1)	 (2)	 (3)	SGMM2R_IS6 (4)	 (5)	 (6)
$\ln(Y/L)_{i,t-1}$	0.929*** (0.018)	0.957*** (0.017)	0.941*** (0.017)	0.949*** (0.017)	0.969*** (0.015)	0.953*** (0.016)
$\ln s_{it}$	0.097*** (0.029)	0.100*** (0.030)	0.088*** (0.029)	0.081*** (0.029)	0.079*** (0.027)	0.076*** (0.028)
$\ln(n_{it} + g + \delta)$	-0.171*** (0.065)	-0.220*** (0.063)	-0.191*** (0.066)	-0.109 (0.078)	-0.143** (0.069)	-0.121 (0.076)
$\ln h_{it}$	0.036 (0.026)	0.023 (0.029)	0.023 (0.025)	0.038 (0.034)	0.030 (0.032)	0.035 (0.032)
SSA dummy	-0.092*** (0.028)		-0.076*** (0.029)	-0.070*** (0.024)		-0.057** (0.026)
EA dummy		0.065** (0.026)	0.044 (0.029)		0.068*** (0.022)	0.045* (0.025)
m_2 (p-value)	0.44	0.45	0.45	0.44	0.44	0.44
Hansen (p-value)	0.23	0.27	0.28	0.24	0.27	0.26
Dif.Hansen (p-value)	0.79	0.83	0.88	0.43	0.47	0.47
No. of instr.$^{(a)}$	123	123	123	111	111	111

Note: *** denotes significance at the one percent level. ** denotes significance at the five percent level. * denotes significance at the ten percent level. Windmeijer-corrected standard errors are reported in parentheses. $(Y/L)_{i,t-1}$ denotes initial output per worker, s_{it} physical investment rate, $(n_{it} + g + \delta)$ growth rate of labour (adjusted by the rate of technological progress and the rate of depreciation) and h_{it} the level of human capital. (a) the number of instruments used for endogenous and predetermined variables. m_2 is Arellano-Bond test of second order serial correlation. The number of observations in all columns is 1170, the number of countries is 134 countries and the average number of periods is 8.7. STATA output of the regressions is reported in appendix II.8.

Table II.4.9: The SGMM2R estimation of the HCASM with SSA and EA dummies – Sample 2

	SGMM2R_IS6		
	(1)	(2)	(3)
$\ln(Y/L)_{i,t-1}$	0.972*** (0.015)	0.987*** (0.013)	0.973*** (0.015)
$\ln s_{it}$	0.072** (0.029)	0.068** (0.028)	0.068** (0.028)
$\ln(n_{it} + g + \delta)$	0.051 (0.054)	0.023 (0.053)	0.044 (0.059)
$\ln h_{it}$	0.044 (0.033)	0.047 (0.035)	0.045 (0.033)
SSA dummy	-0.069*** (0.023)		-0.060** (0.025)
EA dummy		0.051** (0.021)	0.030 (0.022)
m_2 (p-value)	0.43	0.43	0.44
Hansen (p-value)	0.21	0.23	0.25
Dif.Hansen (p-value)	0.35	0.47	0.47
No. of instr.$^{(a)}$	111	111	111

Note: *** denotes significance at the one percent level. ** denotes significance at the five percent level. * denotes significance at the ten percent level. Windmeijer-corrected standard errors are reported in parentheses. $(Y/L)_{i,t-1}$ denotes initial output per worker, s_{it} physical investment rate, $(n_{it} + g + \delta)$ growth rate of labour (adjusted by the rate of technological progress and the rate of depreciation) and h_{it} the level of human capital. (a) the number of instruments used for endogenous and predetermined variables. m_2 is Arellano-Bond test of second order serial correlation. The number of observations in all columns is 1114, the number of countries is 127 countries and the average number of periods is 8.8. STATA output of the regressions is reported in appendix II.9.

Firstly, it can be seen that Arellano-Bond tests in Tables II.4.8 and II.4.9 do not reject the null hypothesis of no second-order serial correlation in the regression models. In Tables II.4.8, the coefficient on the SSA dummy is found to be significant and negative in regressions one and four while that on the EA dummy is found to be significant and positive in regressions two and five. When including both SSA and EA dummies into the HCASM, regressions three and six of Table II.4.8 show the coefficient on the SSA dummy is still significant at conventional levels. However, the coefficient on the EA dummy is insignificant in regression three while weakly significant at the ten percent level in regression six. Similarly, the SGMM2R_IS6 regressions using sample 2 in Table II.4.9 also show that the SSA and EA dummies are significant when they are individually included into the HCASM. When the two dummies are included into the HCASM at the same time, the SSA dummy is found to be significant.

In general, the above results seem to indicate that the HCASM with variables including initial output per worker, the investment rate, labour force growth and human capital does not fully account for the growth of countries in sub-Saharan Africa and East Asia. Alternatively, this study checks this by using a two-step procedure suggested in Hoeffler (2002), by which a GMM estimator is used to estimate the HCASM in the first step to achieve consistent estimates of the coefficients on the regressors, and residuals are then calculated from the estimation. In the second step, we can check directly the correlation between the SSA/EA dummy and the residuals by regressing the residuals on the dummy variables. A finding on the significance of the SSA/EA dummy in the second step indicates that the HCASM cannot fully account for the variation in the growth of SSA/EA countries.

However, different from Hoeffler (2002), this study uses the efficient two-step SGMM with Windmeijer-corrected standard errors, instead of the one-step SGMM, to estimate the HCASM in the first step of the procedure. Moreover, the first step of the procedure is based on a reduced instrument set compared to Hoeffler (2002), as well as using more recent data. In particular, the preferred SGMM2R_IS3 and SGMM2R_IS6 regressions are used for sample 1 while SGMM2R_IS6 regression is used for sample 2 to estimate the HCASM in step one. Results of the two-step procedure using Sample 1 are reported in Table II.4.10 while those using Sample 2 are reported in Table II.4.11. In regressions two, three, six and seven of Table II.4.10 and regressions two and three of Table II.4.11, this study regresses the residuals on the SSA/EA dummies individually while in regressions four and eight of Table II.4.10 and regression four of Table II.4.11, this study regresses the residuals on both the SSA and EA dummies at the same time.

In the two-step procedure using SGMM2R_IS3 in Table II.4.10, the coefficients on the SSA and EA dummies are found to be significant at conventional levels in all the three regressions two, three and four. In the two-step procedure using SGMM2R_IS6 in the same table, the SSA and EA dummies are found to be significant when we regress the estimated residuals on them individually (see regressions six and seven); however, when regressing the estimated residuals on both dummies at the same time, the SSA dummy is weakly significant at the ten percent level while the EA dummy is insignificant (see regression eight in Table II.4.10).

With respect to Table II.4.11 using Sample 2, the SSA and EA dummies are again found to be significant at conventional levels when we regress the residuals (estimated from the SGMM2R_IS6 estimation of the HCASM) on the dummies individually (see column two and three of Table II.4.11). When regressing the estimated residuals on both the SSA and EA dummies simultaneously, the SSA dummy is found to be weakly significant at the ten percent level while the EA dummy is still significant at the five percent level (see column four of Table II.4.11).

It can be seen that the significance of the SSA and EA dummies found in step 2 of most the two-step procedures show that the HCASM does not seem to be able to fully account for the variation in the growth performance experienced by sub-Saharan African and East Asian countries.

As for the SSA dummy particularly, the empirical evidence in Tables II.4.10 and II.4.11 shows that it is significant at conventional levels in all the regressions when we regress the estimated residuals on it alone (see regressions two and five of Table II.4.10 and regression two of Table II.4.11). The results are different from those found in Hoeffler (2002) where she uses a standard one-step system GMM (in step 1 of the two-step procedure) based on a sample of 85 countries for 1960-1990 and establishes the insignificance of the dummy for sub-Saharan African countries. We would like to stress that our study uses two-step SGMM estimations with Windmeijer-corrected standard errors which is suggested to be more efficient in comparison with a standard one-step system GMM (Windmeijer, 2005; Roodman, 2006). In addition, the samples used in our study are larger and more recent than the sample used in Hoeffler (2002). Thus, our findings are making use of additional information.

Table II.4.10: Two-step procedure using SGMM2R_IS3 and SGMM2R_IS6 - Sample 1

	Step 1 SGMM2R_IS3 (1)	Step 2 (2)	(3)	(4)	Step 1 SGMM2R_IS6 (5)	Step 2 (6)	(7)	(8)
$\ln(Y/L)_{i,t-1}$	0.949***				0.969***			
	(0.017)				(0.015)			
$\ln s_{it}$	0.111***				0.088***			
	(0.030)				(0.028)			
$\ln(n_{it}+g+\delta)$	-0.207***				-0.134*			
	(0.062)				(0.071)			
$\ln h_{it}$	0.033				0.030			
	(0.029)				(0.035)			
SSA dummy		-0.049***		-0.042**		-0.040**		-0.031*
		(0.019)		(0.018)		(0.0169)		(0.017)
EA dummy			0.060***	0.050***			0.065***	0.057
			(0.019)	(0.018)			(0.018)	(0.018)

Note: *** denotes significance at the one percent level. ** denotes significance at the five percent level. * denotes significance at the ten percent level. Standard errors that are asymptotically robust to both heteroscedasticity and serial correlation are reported in parentheses. $(Y/L)_{i,t-1}$ denotes initial output per worker, s_{it} investment rate, $(n_{it}+g+\delta)$ growth rate of labour (adjusted by the rate of technological progress and the rate of depreciation) and h_{it} the level of human capital. The number of observations in all columns is 1170, the number of countries is 134 countries and the average number of periods is 8.7. STATA output of the regressions is reported in appendix II.10.

Table II.4.11: Two-step procedure using SGMM2R_IS6 - Sample 2

	Step 1 SGMM2R_IS6 (1)	Step 2 (2)	(3)	(4)
$\ln(Y/L)_{i,t-1}$	0.988***			
	(0.014)			
$\ln s_{it}$	0.075***			
	(0.029)			
$\ln(n_{it} + g + \delta)$	0.026			
	(0.050)			
$\ln h_{it}$	0.042			
	(0.037)			
SSA dummy		-0.037**		-0.031*
		(0.017)		(0.017)
EA dummy			0.048**	0.04**
			(0.0198)	(0.020)

Note: *** denotes significance at the one percent level. ** denotes significance at the five percent level. * denotes significance at the ten percent level. Standard errors that are asymptotically robust to both heteroscedasticity and serial correlation are reported in parentheses. $(Y/L)_{i,t-1}$ denotes initial output per worker, s_{it} investment rate, $(n_{it} + g + \delta)$ growth rate of labour (adjusted by the rate of technological progress and the rate of depreciation) and h_{it} the level of human capital. The number of observations in all columns is 1114, the number of countries is 127 countries and the average number of periods is 8.8. STATA output of the regressions is reported in appendix II.11.

The HCASM and technological progress

The empirical evidence in the previous section indicates that the HCASM does not seem to fully explain the differences in the growth performance of SSA and EA countries. A potential reason explaining the abnormal growth of SSA and EA countries is that rates of technological progress of countries in sub-Saharan Africa and East Asia are different from those of countries in other regions.

It is worth noting that, in the HCASM, the rate of technological progress is assumed to be common to all countries. In this research, as well as in other panel work on the HCASM in the empirical studies, for example, Islam (1995) and CEL, a common but time-varying rate of technological progress is proxied by the period dummies (ω_t) in the regression model. In this sub-section, we further examine whether the differential technological progress of sub-Saharan African and East Asian countries affects their growth of output per worker.

We can do this by firstly creating interaction terms of the SSA and EA dummies with period dummies. In particular, we call interactions of the SSA dummy with period dummies SSAP and call interactions of the EA dummy with period dummies EAP, where

$$SSAP_{rt} = SSA * \omega_t$$

$$EAP_{rt} = EA * \omega_t$$

Next, we use the preferred robust SGMM2R estimators to estimate the HCASM with dummies SSA and EA and interactions SSAP and EAP. The results of the SGMM2R estimation are reported in Table II.4.12 for Sample 1 and reported in Table II.4.13 for Sample 2.

In Table II.4.12 using Sample 1, the regression model in column one and four is the HCASM with the dummy SSA and interactions SSAP, that in column two and five is the HCASM with dummy EA and interactions EAP, and that in column three and six is the HCASM with dummies SSA and EA and interactions SSAP and EAP. Similarly, in Table II.4.13 using Sample 2, the regression model in column one is the HCASM with the dummy SSA and interactions SSAP, that in column two is the HCASM with dummy EA and interactions EAP, and that in column three is the HCASM with dummies SSA and EA and interactions SSAP and EAP. The results of Arellano-Bond tests do not reject the null hypothesis of no second-order serial correlation in the regression models in the two tables.

Firstly, we look at the results of the SGMM2R in columns of Table II.4.12 and compare them with those in the corresponding columns of Tables II.4.8. In columns one and two of Table II.4.12, when we include SSAP or EAP individually into the HCASM with SSA or EA dummy (in table II.4.8), the coefficient on the SSA dummy in the HCASM with SSA and SSAP and that on the EA dummy in the HCASM with EA and EAP in Tables II.4.12 is still significant at conventional levels in the SGMM2R_IS3 estimation. Similarly, the coefficient on the EA dummy remains significant at conventional levels when interactions EAP are included into the HCASM in the SGMM2R_IS6 estimation (see column five of Tables II.4.12 and II.4.8). However, the results in column four of Tables II.4.12 and II.4.8 show that the coefficient on the SSA dummy becomes insignificant when we control for the technological change for SSA countries by introducing interactions SSAP into the regression model. We also find similar empirical evidence for Sample 2: the coefficient on the SSA dummy becomes insignificant when interactions SSAP are included in the HCASM (see column one of Tables II.4.13 and II.4.9) while that on the EA dummy becomes weakly significant at the ten percent level when interactions EAP are introduced into the regression model (see column two of Tables II.4.13 and II.4.9). Furthermore, when both interactions SSAP and EAP are introduced into the regression model to control for the technological progression of countries in sub-Saharan Africa and East Asia, the coefficients on SSA and EA dummies are found not to be significant at the usual critical level in all the results of regressions (see columns three and six of Table II.4.12 and column three of Table II.4.13). Along with that, most results of tests of joint significance of the interactions reject the exclusion of SSAP/EAP from the regression model.

These results may imply that when the difference in the rate of the technological progression among countries in sub-Saharan Africa, East Asian and other regions are taken into account, the HCASM seems to be able to explain the variation in the growth of SSA and EA countries to a greater extent and it may be a factor that helps to explain the contrasting growth performance of SSA and EA countries over the last fifty years.

Table II.4.12: SGMM2R_IS3 and SGMM2R_IS6 estimation of the HCASM with SSA and EA dummies and interactions SSAP and EAP – Sample 1

	SGMM2R_IS3			SGMM2R_IS6		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(Y/L)_{i,t-1}$	0.920*** (0.019)	0.957*** (0.016)	0.944*** (0.023)	0.940*** (0.018)	0.972*** (0.015)	0.947*** (0.024)
$\ln s_{it}$	0.094*** (0.032)	0.094*** (0.027)	0.076*** (0.028)	0.082*** (0.022)	0.083*** (0.027)	0.071*** (0.026)
$\ln(n_{it}+g+\delta)$	-0.142** (0.070)	-0.237*** (0.055)	-0.190*** (0.064)	-0.083 (0.068)	-0.159** (0.066)	-0.105 (0.076)
$\ln h_{it}$	0.051* (0.029)	0.017 (0.029)	0.026 (0.035)	0.056 (0.039)	0.015 (0.037)	0.041 (0.041)
SSA dummy	-0.072** (0.033)		-0.019 (0.035)	-0.036 (0.033)		-0.027 (0.041)
EA dummy		0.088*** (0.033)	0.117* (0.066)		0.076*** (0.028)	0.050 (0.034)
Test joint sign. of SSAP (p-value)	0.08			0.02		
Test joint sign. of EAP (p-value)		0.02			0.01	
Test joint sign. of SSAP and EAP (p-val.)			0.01			0.01
m_2 (p-value)	0.39	0.48	0.43	0.39	0.46	0.39
Hansen (p-value)	0.49	0.63	0.88	0.28	0.53	0.57
Dif.Hansen (p-value)	0.99	0.98	1.00	0.59	0.81	0.98
No. of instr.$^{(a)}$	123	123	123	111	111	111

Note: *** denotes significance at the one percent level. ** denotes significance at the five percent level. * denotes significance at the ten percent level. Windmeijer-corrected standard errors are reported in parentheses. $(Y/L)_{i,t-1}$ denotes initial output per worker, s_{it} investment rate, $(n_{it}+g+\delta)$ growth rate of labour (adjusted by the rate of technological progress and the rate of depreciation) and h_{it} the level of human capital. (a) the number of instruments used for endogenous and predetermined variables. m_2 is the Arellano-Bond test of second order serial correlation. Regressions one, two and three use SGMM2R_IS3. Regressions four, five and six use SGMMR_IS6. The number of observations in all columns is 1170, the number of countries is 134 countries and the average number of periods is 8.7. STATA output of the regressions is reported in appendix II.12.

Table II.4.13: SGMM2R estimation of the HCASM with SSA and EA dummies and interactions SSAY and EAY – Sample 2

	SGMM2R_IS6 (1)	 (2)	 (3)
$\ln(Y/L)_{i,t-1}$	0.965*** (0.016)	0.986*** (0.013)	0.969*** (0.017)
$\ln s_{it}$	0.073** (0.034)	0.073*** (0.027)	0.065** (0.031)
$\ln(n_{it} + g + \delta)$	0.065 (0.061)	0.000 (0.056)	0.054 (0.066)
$\ln h_{it}$	0.055 (0.038)	0.034 (0.037)	0.045 (0.039)
SSA dummy	-0.024 (0.028)		-0.013 (0.033)
EA dummy		0.064* (0.038)	0.062* (0.036)
Test joint sign. of SSAP (p-value)	0.01		
Test joint sign. of EAP (p-value)		0.01	
Test joint sign. of SSAP and EAP (p-val.)			0.01
m_2 (p-value)	0.38	0.44	0.45
Hansen (p-value)	0.55	0.45	0.72
Dif.Hansen (p-value)	0.73	0.85	0.97
No. of instr.$^{(a)}$	111	111	111

Note: *** denotes significance at the one percent level. ** denotes significance at the five percent level. * denotes significance at the ten percent level. Windmeijer-corrected standard errors are reported in parentheses. $(Y/L)_{i,t-1}$ denotes initial output per worker, s_{it} investment rate, $(n_{it} + g + \delta)$ growth rate of labour (adjusted by the rate of technological progress and the rate of depreciation) and h_{it} the level of human capital. (a) the number of instruments used for endogenous and predetermined variables. m_2 is the Arellano-Bond test of second order serial correlation. Regressions one, two and three use the SGMM2R_IS6. The number of observations in all columns is 1114, the number of countries is 127 countries and the average number of periods is 8.8. STATA output of the regressions is reported in appendix II.13.

Above, this chapter reported empirical results of the estimation of the HCASM. Empirical results of the investigation on whether the HCASM can account for the growth of sub-Saharan African and East Asian countries were also reported in this chapter. The next chapter will provide the conclusion of this study.

CHAPTER II.5:

CONCLUSION

This study uses the newest version of the Penn World Table, version 8.0, recently published by Feenstra, Inklaar and Timmer (2013) and the newest Educational Attainment Dataset version 1.3 recently published by Barro and Lee (2013) to estimate the human-capital augmented Solow model and test whether the human-capital augmented Solow model can account for the growth of countries in sub-Saharan Africa and East Asia – two regions which experienced contrasting growth performances over the past 50 years.

Using the efficient two-step system GMM estimator with Windmeijer-corrected standard errors to estimate the human-capital augmented Solow model and checking robustness to reductions in the instrument count, as suggested in Roodman (2009), this study provides empirical evidence confirming the conditional-convergence hypothesis of the human-capital augmented Solow model. The rate of the convergence found in this study ranges from 0.3 to roughly 1 per cent a year, which is slower than found in research by Mankiw, Romer and Weil (1992) ($\lambda = 1.4\,\%$), Islam (1995) ($\lambda = 3.8\%$), Caselli, Esquivel and Lefort (1996) ($\lambda = 6.8\%$), Bond, Hoeffler and Temple (2001) ($\lambda = 1.7\%$) and Hoeffler (2002) ($\lambda = 3.3\%$). It is closest to the studies which have used system GMM as their main approach. The estimated rate of the convergence implies that the average time an economy takes to cover half of the distance between its position and its steady state[13] ranges from 69 to 230 years. Empirical evidence shows that the coefficient on the investment rate in physical capital is robustly significant and positive in regressions. This stresses the important role of the investment rate in physical capital towards influencing the level of the growth path as the Solow model

---

[13] See footnote 9 on p.74.

predicts. The influence of the level of human capital on the level of the growth path is found to be insignificant while that of labour-force growth is found not to be robust.

In the second section, this study tests whether the human-capital augmented Solow model can account for the differential growth of countries in sub-Saharan Africa and East Asia by adding regional (SSA and/or EA) dummies into the human-capital Solow model and using a two-step procedure as in Hoeffler (2002). Using two-step system GMM estimation with Windmeijer-corrected standard errors to address country-specific effects and endogeneity, we find that SSA and/or EA dummies are likely to be significant in most cases. This implies that the human-capital capital Solow model seems to be unable to fully account for the variation in unusual growth performance of sub-Saharan and East Asian countries even when country-specific effects and endogeneity are taken into account.

As for the SSA dummy, we find that it is significant in all cases when it is tested individually. This evidence is different from that in the research by Hoeffler (2002) who uses a standard one-step system GMM estimator based on a sample of 85 countries for 1960-1990. We would like to stress that our study uses the two-step SGMM estimations with Windmeijer-corrected standard errors which is suggested to be more efficient in comparison with a standard one-step system GMM (Windmeijer, 2005; Roodman, 2006). In addition, the samples used in our study are larger and the PWT data are more recent than the data and sample used in Hoeffler (2002). Thus, our findings are making use of additional information.

This study further allows for differences in the rate of the technological progress of SSA and EA countries by introducing interactions (SSAP/EAP) between the SSA/EA dummies and the period dummies into the regression model, to consider the effect of technological progress on the growth of SSA and EA countries. We find that when both interactions SSAP and EAP are introduced into the regression model to control for the technological progress of countries in sub-Saharan Africa and East Asia, the coefficients on the SSA and EA dummies are found not to be significant at the usual critical levels in all the regressions. In addition, the results of tests of joint significance of the interactions reject the exclusion of SSAP and EAP from the regression models. This implies the rates of technological progress between the two regions are likely to be different and this may be a reason that helps to explain the contrasting growth performance experienced by sub-Saharan African and East Asian countries over the past fifty years.

Finally, it can be seen that one contribution of this study is the use of the newest Penn World Table version 8.0 recently published by Feenstra, Inklaar and Timmer (2013) and the newest Educational Attainment Dataset version 1.3 recently published by Barro and Lee (2013) in the research. Therefore, this study's findings are making use of additional information. The second contribution is that the estimation of the human-capital Solow model is much closer in line with current best practice, as set out by Windmeijer (2005) and Roodman (2006, 2009), when compared to the earlier literature, which was not. For example earlier papers, such as Bond, Hoeffler and Temple (2001) or Hoeffler (2002), did not restrict the instrument set, nor use efficient two-step system GMM with Windmeijer-corrected standard errors. Given what is now known about the performance of these estimators in such circumstances, the approach taken in this study is more likely to be informative.

REFERENCES

Agarwal, J.P. (1980) Determinants of foreign direct investment: A survey. *Weltwirtschaftliches Archiv*. Vol. 116(4), pp. 739-773.

Aliber, R.Z. (1970) A theory of direct foreign investment. In: C. P. Kindleberger, ed., *The international corporation*. Massachusetts: MIT Press, pp. 17-34.

Baltagi, B.H. (2008) *Econometric analysis of panel data*. 4th ed. Chichester: John Wiley and Sons.

Barro, R.J. (1991) Economic growth in a cross section of countries. *Quarterly Journal of Economics*. Vol. 106(2), pp. 407-443.

Barro, R.J. and Lee, J. (2013) A new set of educational attainment in the world, 1950-2010. *Journal of Development Economics*. Vol.104, pp. 184-198.

Barro, R.J. and Sala-i-Martin, X. (2004) *Economic growth*. 2nd ed. Cambridge, MA: The MIT Press.

Benassy-Quere, A., Fontagne, L., and Lahreche-Revil, A. (2001) Exchange-rate strategies in the competition for attracting foreign direct investment. *Journal of the Japanese and International Economics*. Vol. 15, pp, 178-198.

Bennet, P.D., and Green, R.T. (1972) Political instability as a determinant of direct foreign investment in marketing. *Journal of Marketing Research*. Vol. 9(2), pp. 182-186.

Bevan, A.A. and Estrin, S. (2004) The determinants of foreign direct investment into European transition economies. *Journal of Comparative Economics*. Vol. 32(4), pp. 775–787.

Blundell, R. and Bond, S. (1998) Initial conditions and moment restrictions in dynamic panel data model. *Journal of Econometrics*. Vol. 87, pp. 115-143.

Bond, S. (2002) Dynamic panel data models: A guide to micro data methods and practice. Cemmap Working Paper CWP09/02.

Bond, S., Hoeffler, A. and Temple, J. (2001) GMM estimation of empirical growth models. CEPR Discussion Paper 3048.

Bruno, G.S.F. (2005) Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models. *Economics Letters*. Vol. 87, pp. 361-366.

Buckley, P.J. and Casson, M. (1976) *The future of multinational enterprise*. London: MacMillan.

Bun, M.J.G., and Kiviet, J.F. (2003) On the diminishing returns of higher order terms in asymptotic expansions of bias. *Economics Letters*. Vol. 79, pp. 145 – 152.

Carroll, C.D. and Weil, D.N. (1994) Savings and growth: A reinterpretation. *Carnegie-Rochester Conference Series on Public Policy*. Vol. 40, pp. 133-192.

Caselli, F., Esquivel, G., and Lefort, F. (1996) Reopen the convergence debate: A new look at cross-country growth empirics. *Journal of Economic Growth*. Vol. 1(3), pp. 363-389.

Chudik, A. and Pesaran, M.H (2013) Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. CESifo WP Number 4232.

Clegg, J. (1987) *Multinational Enterprise and World Competition*. London: MacMillan.

Coakley, J., Fuertes, A., and Smith, R. (2006) Unobserved heterogeneity in panel time series model. *Computational Statistics and Data Analysis*. Vol. 50, pp. 2361-2380.

Coase, R.H. (1937) The nature of the firm. *Economica*. Vol. 4(16), pp. 386-405.

Crowley, P. and Lee, J. (2003) Exchange rate volatility and foreign investment: International evidence. *The International Trade Journal*. Vol. 17 (3), pp. 227-252.

Cushman, D.O. (1988) Real exchange rate risk, expectations and the level of direct investment. *Review of Economics Statistics*. Vol. 67(2), pp. 297-308.

Davidson, W.H. (1980) The location of foreign direct investment activity: Countries characteristics and experience effects. *Journal of International Business Studies*. Vol. 11(2), pp. 9-22.

Duffy, J. and Papageorgiou, C. (2000) A cross-country empirical investigation of the aggregate production function specification. *Journal of Economic Growth*. Vol. 5, pp. 87-120.

Dunning, J.H. (1981) *International production and the multinational enterprise*. London: Allen and Unwin.

Dunning, J.H. (1988) *Explaining international production*. London: Unwin Hyman.

Dunning, J.H. (1992) *Multinational enterprises and the global economy*. Wokingham: Addison-Wesley.

Durlauf, S.N., Johnson, P.A., and Temple, J.R.W. (2005) Growth econometrics. In: Aghion, P. and Durlauf, S.N., eds., *Handbook of Economic Growth: Volume 1A*. Amsterdam: North-Holland, 2005, pp. 555-677.

Easterly, W. and Levine, R. (1998) Troubles with the neighbours: Africa's problems, Africa's Opportunity. *Journals of African Economies*. Vol. 7(1), pp. 120-142.

Eberhardt, M. and Bond, S. (2009) Cross-section dependence in nonstationary panel models: A novel estimator. MPRA Paper No. 17870.

Eberhardt, M. (2011) Panel time-series modeling: New tools for analyzing xt data. UK Stata Users Group Meeting, Cass Business School, London.

Feenstra, R.C., Inklaar, R. and Timmer, M.P. (2013) *The next generation of the Penn World Table*. Available for download at www.ggdc.net/pwt

Fisher, R.A. (1932) *Statistical Methods for Research Workers*. 4th ed. Edinburgh: Oliver and Boyd.

Giddy, I.H. (1978) The demise of the product cycle model in international business theory. *Columbia Journal of World Business*. Vol.13, pp. 90-97.

Goldberg, L.S. and Kolstad, C.D. (1995) Foreign direct investment, exchange rate variability and demand uncertainty. *International Economic Review*. Vol. 36(4), pp. 855-873.

Gorg, H. and Wakelin, K. (2002) The impact of exchange rate volatility on US direct investment. *The Manchester School*. Vol. 70(3), pp. 380-397.

Hall, A., 1990. *Testing for a unit root in time series with pretest data-based model selection*. North Carolina State University, Working Paper.

Helpman, E. (1984) A simple theory of international trade with multinational corporate. *Journal of Political Economy*. Vol. 92(3), pp. 451-471.

Hoeffler, A. (1998) Econometric studies of growth, convergence and conflicts. PhD Thesis. University of Oxford.

Hoeffler, A. (2002) The augmented Solow model and the African growth debate. *Oxford Bulletin of Economics and Statistics*. Vol. 64(2), pp. 135-158.

Hsiao, C. (1986) *Analysis of panel data*. Cambridge, MA: Cambridge University Press

Hymer, S.H. (1976) *The international operations of national firms: A study of foreign direct investment*. Cambridge: MIT Press.

Inada, K. (1963) On a two-sector model of economic growth: Comment and a generalisation. *Review of Economics Studies*. Vol. 30(2), pp. 119-127.

Islam, N. (1995) Growth empirics: A panel data approach. *Quarterly Journal of Economics*. Vol. 110(4), pp. 1127-1170.

Itagaki, T. (1981) The theory of the multinational firm under exchange rate uncertainty. *The Canadian Journal of Economics*. Vol. 14(2), pp. 276-297.

Iversen, C. (1935) *Aspects of international capital movements*. London: Levin and Munksgaard.

Janicki, H.P., and Wunnava, P.V. (2004) Determinants of foreign direct investment: Empirical evidence from EU accession candidates. *Applied Economics*. Vol. 36, pp. 505-509.

Judson, R.A., and Owen, A.L. (1999) Estimating dynamic panel data models: a guide for macroeconomists. *Economics Letters*. Vol. 65, pp. 9-15.

Kapetanios, G., Pesaran, M.H., and Yamagata, T. (2011) Panels with non-stationary multifactor error structures. *Journal of Econometrics*. Vol. 160(2), pp. 326-348.

Kindleberger, C.P. (1969) *American business abroad: Six lectures on direct investment*. New Heaven: Yale University Press.

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

Kiviet, J.F. (1999) Expectation of expansions for estimators in a dynamic panel data model: some results for weakly exogenous regressors. In: Hsiao, C., Lahiri, K., Lee, L.F., Pesaran, M.H., eds., *Analysis of Panel Data and Limited Dependent Variables*. Cambridge: Cambridge University Press, pp. 199-225.

Knickerbocker, F.T. (1973) *Oligopolistic reaction and multinational enterprise*. Boston: Harvard University Press.

Koechlin, T. (1992) The determinants of the location of USA direct foreign investment. *International Review of Applied Economics*. Vol. 6(2), pp. 203-216.

Kohlhagen, S.W. (1977) The effect if exchange-rate adjustments on international investment: Comment. In: P.B. Clark, D.E. Logue and R.Sweeney, eds., *The effect of exchange rate adjustments*. Washington, DC: US Government Printing Office, pp. 194-197.

Kojima, K. (1977) *Japan and a new world economic order*. London: Croom Helm.

Lee, J.Y. and Mansfield, E. (1996) Intellectual property protection and U.S. foreign direct investment. *Review of Economics and Statistics*. Vol. 78(2), pp. 181-186.

Lessard, D.R. (1976) World, country and industry relation in equity returns: Implication for risk reduction through international diversification. *Financial Analysts Journal*. Vol. 32(1), pp. 32-38.

Levin, A., and Lin, C. (1992) *Unit root tests in panel data: Asymptotic and finite-sample properties*. UC San Diego, Working Paper 92-23.

Levin, A., Lin, C. and Chu, C. (2002) Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics*. Vol. 108(1), pp. 1-24.

Levine, R. and Renelt, D. (1992) A sensitivity analysis of cross-country growth regressions. *American Economic Review*. Vol. 82(4), pp. 942-963.

Limao, N. and Venables, A.J. (2001) Infrastructure, geographical disadvantage and transport costs. *World Bank Economic Review*. Vol.15(3), pp. 451-479.

Loree, D.W. and Guisinger S.E. (1995) Policy and non-policy determinants of U.S. equity foreign direct investment. *Journal of International Business Studies*. Vol. 26 (2), pp. 281-299.

Lucas, R.E. (1988) On the mechanics of economic development. *Journal of Monetary Economics*. Vol. 22(1), pp. 3-42.

Maddala, G.S. and Wu, S. (1999) A comparative study of unit root tests with panel data a new test. *Oxford Bulletin of Economics and Statistics*. Vol. 61, pp. 631-652.

Mankiw, N.G., Romer, D., and Weil, D.N. (1992) A contribution to the empirics of economic growth. *Quarterly Journal of Economics*. Vol. 107(2), pp. 407-437.

Markowitz, H. (1959) *Portfolio analysis*. New York: Wiley

Markusen, J.R. (1984) Multinationals, multi-plant economies and the gains from trade. *Journal of International Economics*. Vol. 16(3-4), pp. 205-226.

Masanjala, W.H. and Papageorgiou, C. (2004) The Solow model with CES technology: Nonlinearities and parameter heterogeneity. *Journal of Applied Econometrics*. Vol. 19(2), pp. 171-201.

Moosa, I.A. (2002) *Foreign direct investment: Theory, evidence and practice*. New York: Palgrave.

Moosa, I.A. and Cardak, B.A. (2006) The determinants of foreign direct investment: An extreme bound analysis. *Journal of Multinational Financial Management*. Vol. 16, pp. 199-211.

Narula, R. (1996) *Multinational investment and economic structure: Globalization and competitiveness*. London: Routledge.

Nigh, D. (1985) The effects of political events on United States direct foreign investment: A pooled series cross-sectional analysis. *Journal of International Business Studies*. Vol. 16(1), pp. 1-17.

Noorbakhsh, F., Paloni, A. and Yousseff, A. (2001) Human capital and FDI inflows to developing countries: New empirical evidence. *World Development*. Vol. 29(9), pp. 1593–1610.

Parry, T.G. (1985) Internalisation as a general theory of foreign direct investment: A critique. *Weltwirtschaftliches Archiv*. Vol.121(3), pp. 564-569.

Pesaran, M.H. (2004) *General Diagnostic Tests for Cross Section Dependence in Panels*. IZA Discussion Paper No. 1240.

Pesaran, M.H. (2006) Estimation and inference in large heterogeneous panels with multifactor error structure. *Econometrica*. Vol. 74(4), pp.326-348.

Pesaran, M.H. (2007) A simple panel unit roots in the presence of cross-section dependence. *Journal of Applied Econometrics*. Vol. 22, pp. 265-312.

Pesaran, M.H. and Smith, R.P. (1995) Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*. Vol. 68(1), pp. 79-113.

Pesaran, M.H. and Tosetti, E. (2011) Large panels with common factors and spatial correlation. *Journal of Econometrics*. Vol. 161, pp. 182-202.

Petrochilos, G.A. (1989) *Foreign direct investment and the development process: The case of Greece*. Avebury: Gower.

Pritchett, L. (1996) Where has all the education gone? World Bank, Policy Research Working Paper, 1681.

Romer, D. (2006) *Advanced macroeconomics*. 3rd ed. New York: McGraw-Hill.

Romer, P.M. (1986) Increasing returns and long-run growth. *Journal of Political Economy*. Vol. 94, pp. 1002-1037.

Rugman, A.M. (1976) Risk reduction by international diversification. *Journal of International Business*. Vol. 7(2), pp. 75-80.

Rugman, A.M. (1986) New theories of the multinational enterprise: An assessment of internalisation theory. *The Bulletin of Economic Research*. Vol.3 8(2), pp. 101-118.

Sachs, J.D and Warner, A.M. (1997) Sources of slow growth in African Economies. *Journal of African Economies*. Vol. 6(3), pp. 335-376.

Scaperlanda, A.E. and Mauer, L.J. (1969) The determinants of U.S. direct investment in the E.E.C. *American Economic Review*. Vol. 59(4), pp. 558-568.

Schmitz, A. and Beiri, J. (1972) EEC tarriffs and U.S. direct investment. *European Economic Review*. Vol. 3, pp. 259-270.

Schneider, F. and Frey, B. (1985) Economic and political determinants of foreign direct investment. *World Development*. Vol. 13(2), pp. 161-175.

Solomon, L.D. (1978) *Multinational corporations and the emerging world order*. New York: Kennikat Press.

Solow, R.M. (1956) A contribution to the theory of economic growth. Quarterly Journal of Economics. Vol. 70(1), pp. 65-94.

Stock, J.H. and Watson, M.W. (2008) Forecasting in dynamic factor models subject to structural instability. In: J. Castle and N. Shephard, eds., *The methodology and practice of econometrics, A festschrift in honour of Professor David F. Hendry*. Oxford: Oxford University Press.

Summers, R. and Heston, A. (1988) A new set of international comparisons of real product and price levels estimates for 130 countries, 1950-1985. *Review of In-come and Wealth*.Vol.34, pp. 1-25.

Swan, T.W. (1956) Economic growth and capital accumulation. *Economic Record*. Vol. 32(2), pp. 334 -361.

Swenson, D.L. (1994) The impact of U.S. tax reform on foreign direct investment in the United States. *Journal of Public Economics*. Vol. 54(2), pp. 243-266.

Temple, J. (1998) Equipment investment and the Solow model. *Oxford Economic Papers*. Vol. 50, pp. 39-62.

Temple, J. (1999) The new growth evidence. *Journal of Economic Literature*. Vol. 37(1), pp. 112-156.

Topel, R. (1999) Labor markets and economic growth. In: Ashenfelter, A. and Card, D., eds., (1999) *Handbooks in labor economics, Volume 3*. Amsterdam, New York and Oxford: Elsevier Science, North-Holland, pp. 2943-2984.

UNCTAD (2006) Efficient transport and trade facilitation to improve participation by developing countries in international trade. TD/B/COM.3/80.

Vernon, R. (1966) International Investment and International Trade in the Product Cycle. *Quarterly Journal of Economics*. Vol. 80, pp. 190-207.

Vernon, R. (1971) *Sovereignty at bay: The multinational spread of U.S. enterprises*. London: Pelican.

Wei, S. (2000) How taxing corruption on international investors? *Review of Economics and Statistics*. Vol. 82(1), pp. 1-11.

Wheeler, D. and Mody, A. (1992) International investment location decisions: The case of U.S. firms. *Journal of International Economics*. Vol. 33 (1-2), pp. 57–76.

Yang, J., Groenewold, N. and Tcha, M. (2000) The determinants of foreign direct investment in Australia. *The Economic Record*. Vol. 76(232), pp. 45-54.

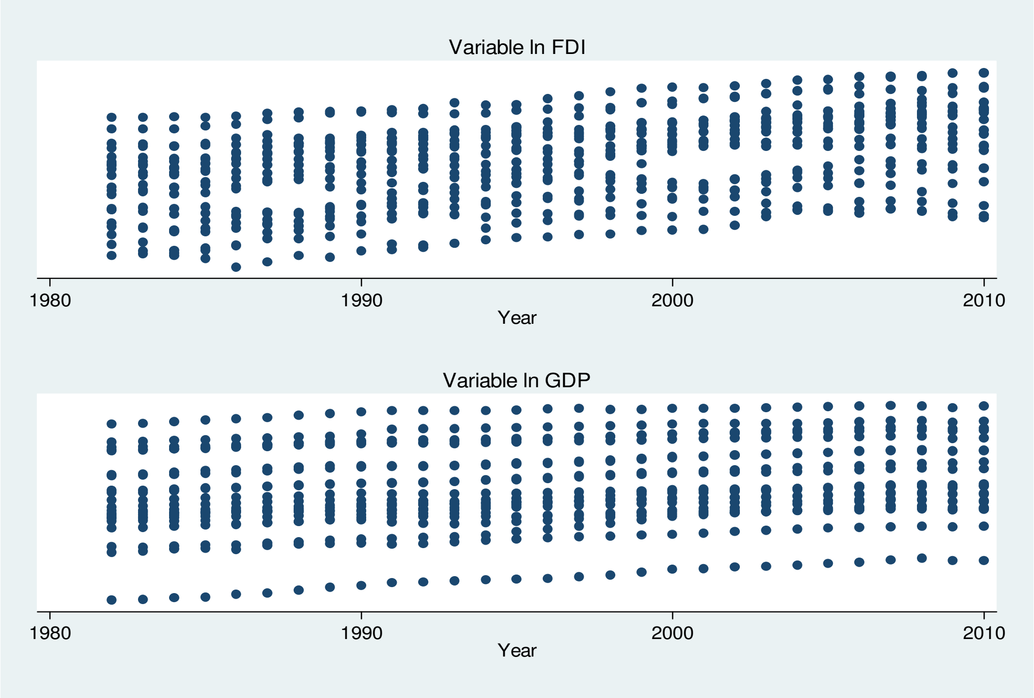
APPENDICES OF STUDY I

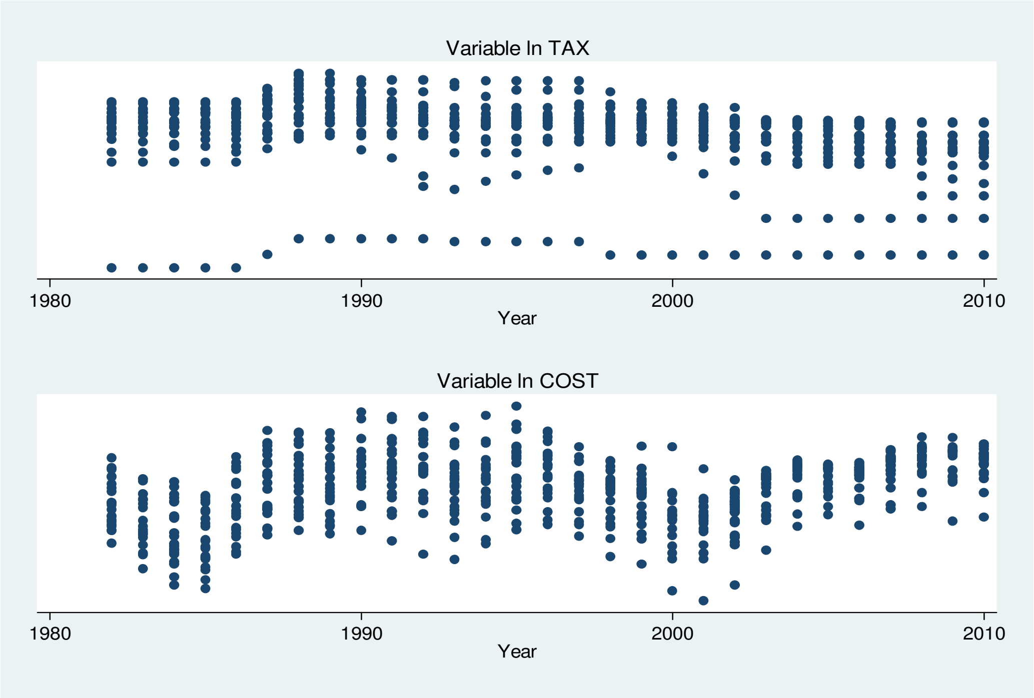
Appendix I.1:

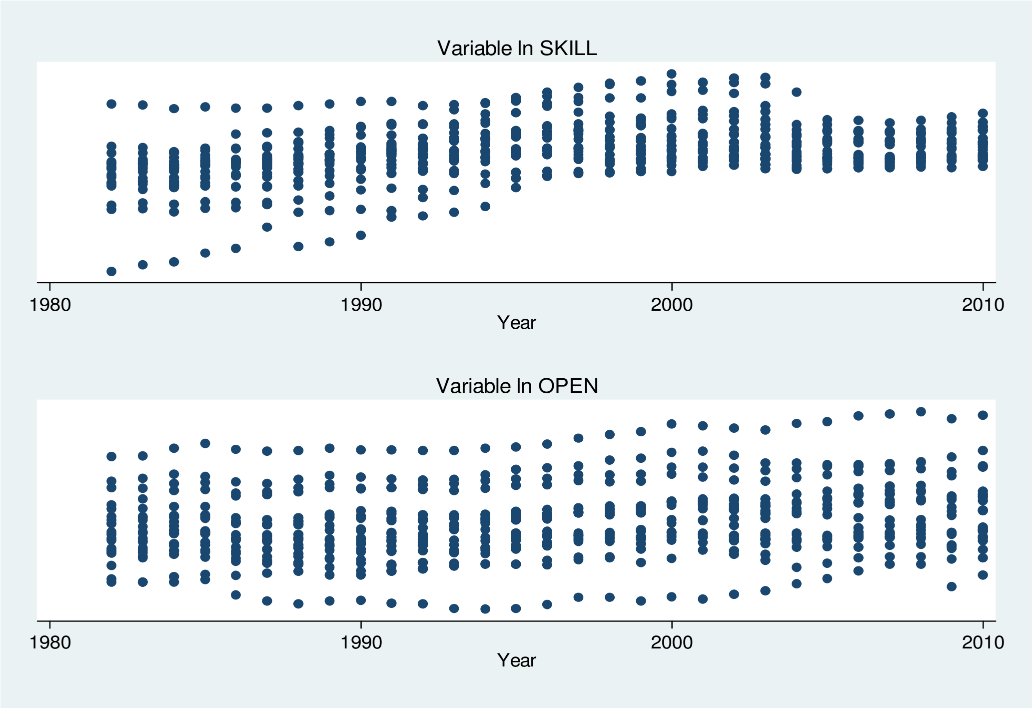
List of countries in the sample

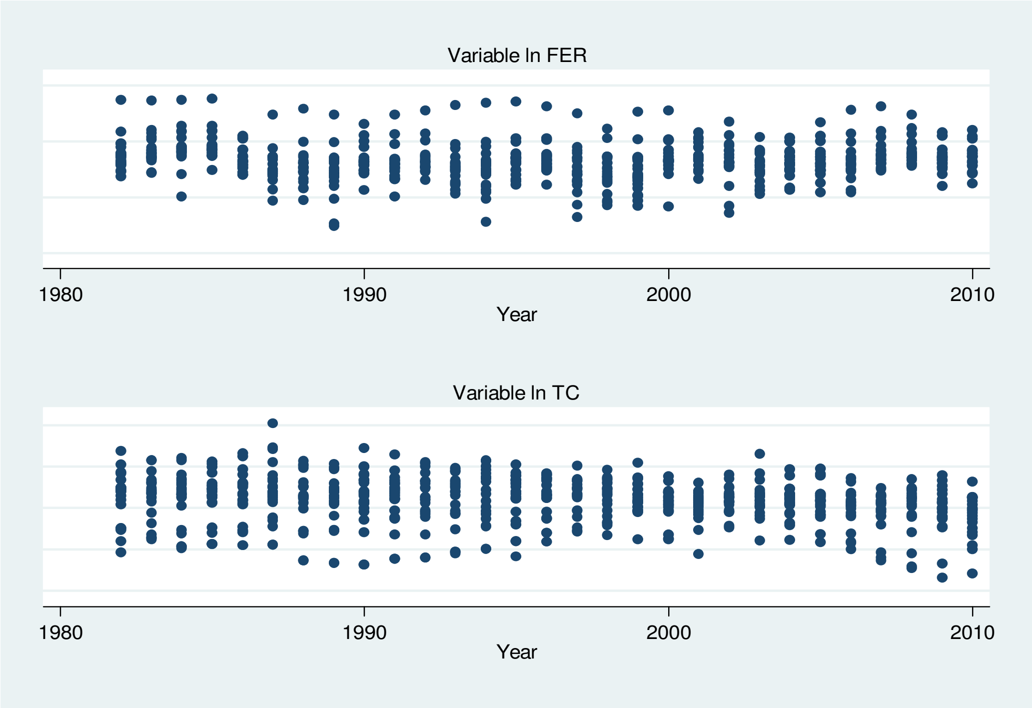
Australia	Germany	New Zealand
Austria	Greece	Norway
Belgium	Ireland	Portugal
Canada	Italy	Spain
Demark	Japan	Sweden
Finland	Luxembourg	Switzerland
France	Netherlands	United Kingdom

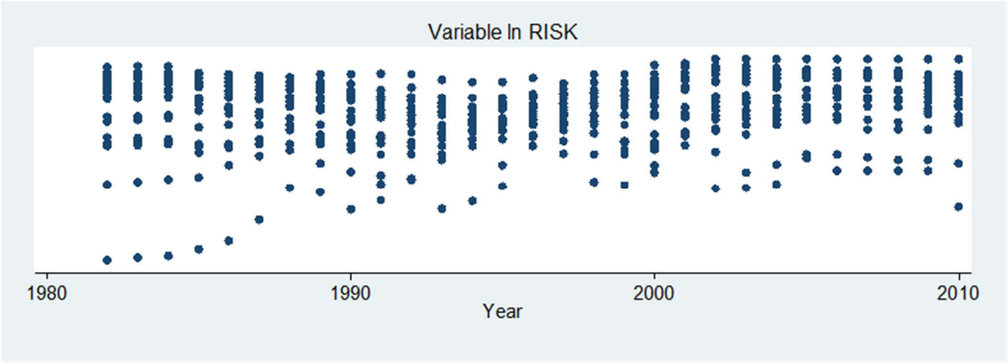
Appendix I.2: Plot of variables over time





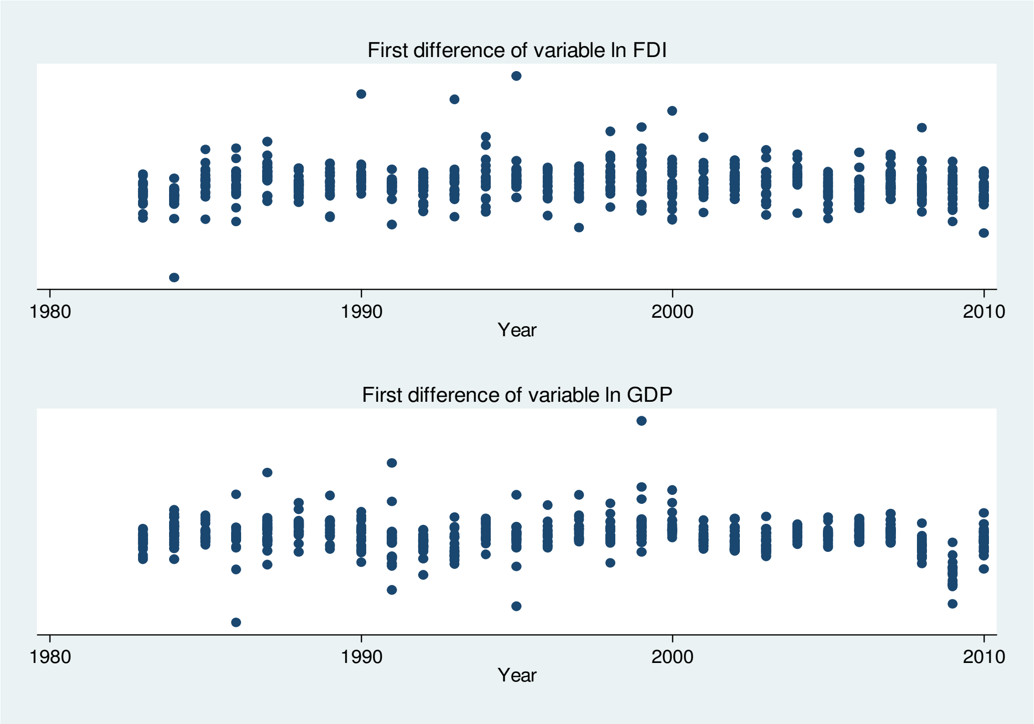

Variable ln SKILL
1980
1990
2000
2010
Year
Variable ln OPEN
1980
1990
2000
2010
Year

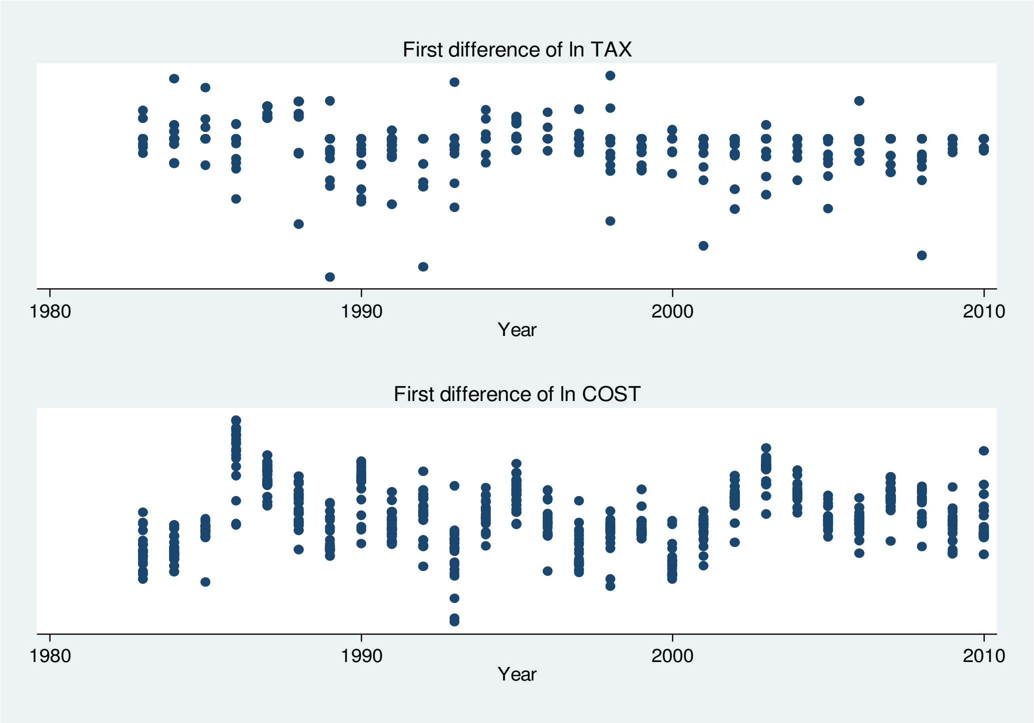

Variable ln FER
1980
1990
2000
2010
Year
Variable ln TC
1980
1990
2000
2010
Year

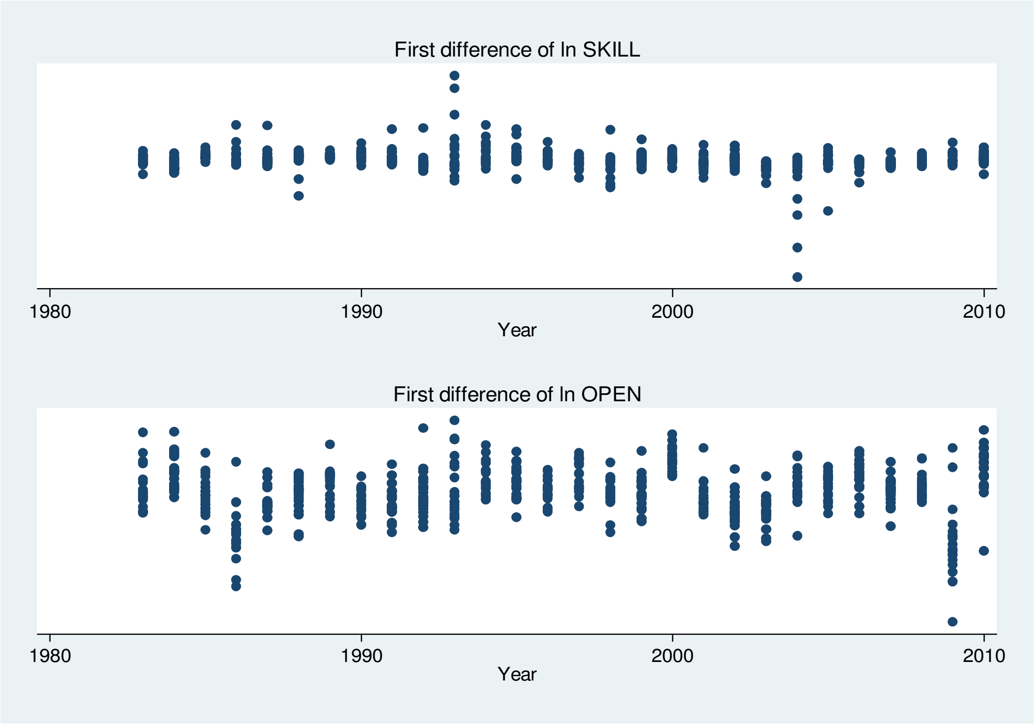


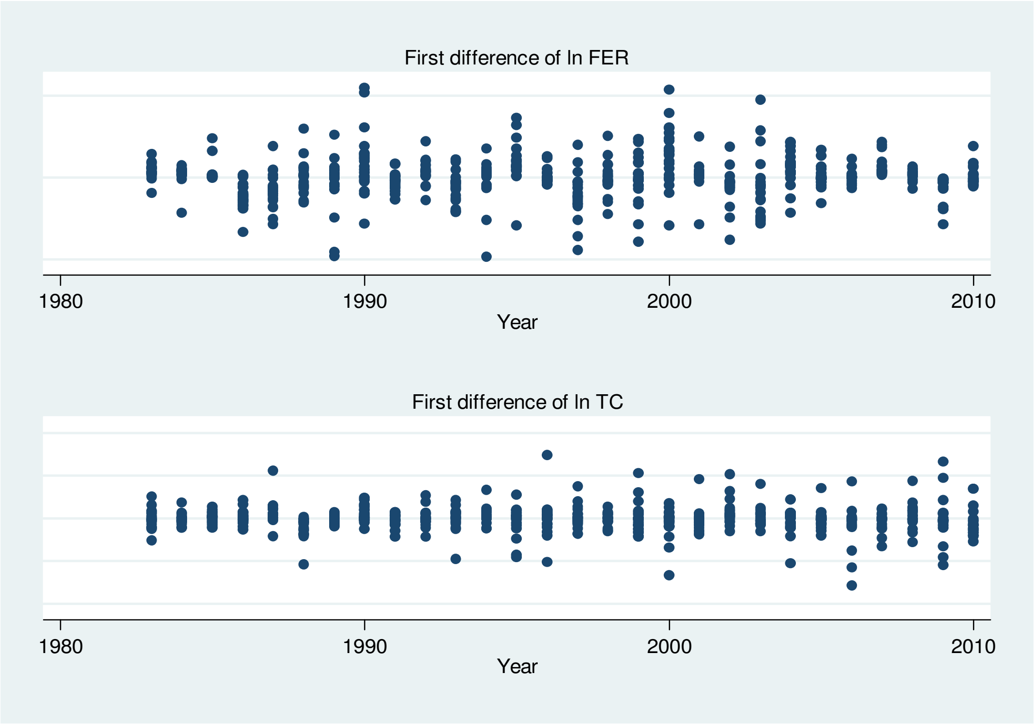
Appendix I.3:

Plot of first difference of variables over time

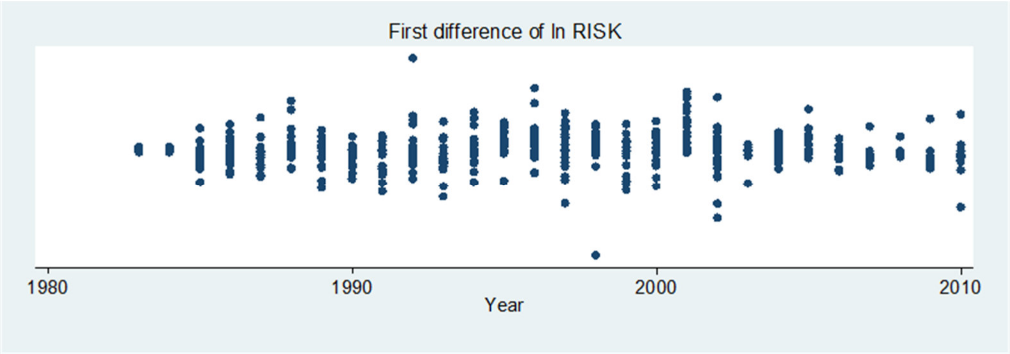



First difference of ln TAX
1980
1990
2000
2010
Year
First difference of ln COST
1980
1990
2000
2010
Year


First difference of ln SKILL
1980
1990
2000
2010
Year
First difference of ln OPEN
1980
1990
2000
2010
Year


First difference of ln FER

First difference of ln TC


First difference of ln RISK

Appendix I.4:

STATA output of regressions of the models using one-year lagged values for independent variables

```
//Pooled OLS
i.year            _Iyear_1982-2010    (naturally coded; _Iyear_1982 omitted)
note: _Iyear_2009 omitted because of collinearity
Linear regression                                      Number of obs =     588
                                                       F( 19,    20) =       .
                                                       Prob > F      =       .
                                                       R-squared     =  0.7589
                                                       Root MSE      =  .86988
                                    (Std. Err. adjusted for 21 clusters in id)
------------------------------------------------------------------------------
             |               Robust
       lnFDI |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     lnGDPL1 |    1.44639   .1699709     8.51   0.000     1.091837    1.800943
     lnTAXL1 |  -.7245821   .1703599    -4.25   0.000    -1.079947   -.3692176
    lnCOSTL1 |  -1.978785   .7170996    -2.76   0.012    -3.474628   -.4829412
   lnSKILLL1 |   1.084278   1.155997     0.94   0.359     -1.32709    3.495646
    lnOPENL1 |   1.662067   .5233484     3.18   0.005     .5703809    2.753752
     lnFERL1 |  -.0612462   .0906707    -0.68   0.507    -.2503819    .1278894
      lnTCL1 |   -.579938   .4430588    -1.31   0.205    -1.504142    .3442664
    lnRISKL1 |   1.838127   1.140266     1.61   0.123    -.5404263     4.21668
 _Iyear_1983 |  -.5086486   .3055888    -1.66   0.112    -1.146096    .1287986
 _Iyear_1984 |  -.8575978   .3397246    -2.52   0.020    -1.566251   -.1489446
 _Iyear_1985 |  -1.100724   .3753274    -2.93   0.008    -1.883643   -.3178049
 _Iyear_1986 |  -1.109149   .3827567    -2.90   0.009    -1.907565   -.3107321
 _Iyear_1987 |  -.4523621   .3286724    -1.38   0.184    -1.137961    .2332365
 _Iyear_1988 |   -.064221   .3524549    -0.18   0.857     -.799429     .670987
 _Iyear_1989 |   .0090001   .3684549     0.02   0.981    -.7595833    .7775835
 _Iyear_1990 |  -.1115786   .3594869    -0.31   0.759    -.8614551    .6382979
 _Iyear_1991 |   .1734601   .3120333     0.56   0.584    -.4774299    .8243502
 _Iyear_1992 |   .1265151   .3451705     0.37   0.718     -.593498    .8465281
 _Iyear_1993 |   .1972271   .3123038     0.63   0.535    -.4542273    .8486815
 _Iyear_1994 |  -.0469462   .3622956    -0.13   0.898    -.8026816    .7087891
 _Iyear_1995 |   .0125299   .3769833     0.03   0.974    -.7738435    .7989033
 _Iyear_1996 |   .1430173   .3133828     0.46   0.653    -.5106877    .7967223
 _Iyear_1997 |   .0252312   .3056737     0.08   0.935    -.6123929    .6628553
 _Iyear_1998 |  -.2387615   .3556266    -0.67   0.510    -.9805855    .5030625
 _Iyear_1999 |  -.2633391   .3688846    -0.71   0.484    -1.032819    .5061407
 _Iyear_2000 |  -.3636693   .3969089    -0.92   0.370    -1.191607    .4642683
 _Iyear_2001 |  -.7646149   .3906714    -1.96   0.064    -1.579541    .0503114
 _Iyear_2002 |  -.8824644    .396265    -2.23   0.038    -1.709059   -.0558701
 _Iyear_2003 |  -.6704537    .387297    -1.73   0.099    -1.478341    .1374336
 _Iyear_2004 |  -.1817926   .3371097    -0.54   0.596    -.8849911    .5214059
 _Iyear_2005 |  -.0382072       .204    -0.19   0.853    -.4637438    .3873293
 _Iyear_2006 |  -.1241344    .185993    -0.67   0.512     -.512109    .2638402
 _Iyear_2007 |  -.2028052    .134699    -1.51   0.148    -.4837823     .078172
 _Iyear_2008 |  -.0642716   .0511956    -1.26   0.224    -.1710637    .0425205
_ Iyear_2009 |   (omitted)
 _Iyear_2010 |   .1551716    .162653     0.95   0.351    -.1841165    .4944598
       _cons |  -13.40555   3.989855    -3.36   0.003    -21.72824   -5.082855
------------------------------------------------------------------------------
```

```
//Fixed Effects
i.year            _Iyear_1982-2010    (naturally coded; _Iyear_1982 omitted)
note: _Iyear_2008 omitted because of collinearity
Fixed-effects (within) regression               Number of obs      =       588
Group variable: id                              Number of groups   =        21
R-sq:  within  = 0.8313                         Obs per group: min =        28
       between = 0.4754                                        avg =      28.0
       overall = 0.5110                                        max =        28
                                                F(20,20)           =         .
corr(u_i, Xb)  = -0.6773                        Prob > F           =         .
                                    (Std. Err. adjusted for 21 clusters in id)
------------------------------------------------------------------------------
             |               Robust
       lnFDI |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     lnGDPL1 |   2.009698   .6911931     2.91   0.009     .5678949    3.451502
     lnTAXL1 |  -.1065776   .2293512    -0.46   0.647    -.5849959    .3718407
    lnCOSTL1 |   .7312111   .4124151     1.77   0.091    -.1290716    1.591494
   lnSKILLL1 |   .7542806    .588609     1.28   0.215    -.4735362    1.982097
    lnOPENL1 |   1.199265   .5813914     2.06   0.052    -.0134966    2.412026
     lnFERL1 |  -.0479594   .0193857    -2.47   0.022    -.0883974   -.0075215
      lnTCL1 |   .3192905   .2720883     1.17   0.254    -.2482757    .8868567
    lnRISKL1 |   .0413476   .4340327     0.10   0.925    -.8640287    .9467239
 _Iyear_1983 |   .0024642   .6377919     0.00   0.997    -1.327946    1.332875
 _Iyear_1984 |   -.063644   .6099233    -0.10   0.918    -1.335922    1.208634
 _Iyear_1985 |  -.0243989   .6052972    -0.04   0.968    -1.287027    1.238229
 _Iyear_1986 |   .0299563   .5864003     0.05   0.960    -1.193253    1.253166
 _Iyear_1987 |   .0523779   .5405796     0.10   0.924    -1.075251    1.180007
 _Iyear_1988 |  -.1068429    .475782    -0.22   0.825    -1.099307    .8856209
 _Iyear_1989 |  -.1087719   .4548848    -0.24   0.813    -1.057645    .8401011
 _Iyear_1990 |  -.0498176   .4383274    -0.11   0.911    -.9641526    .8645173
 _Iyear_1991 |  -.1315521   .3918467    -0.34   0.741    -.9489299    .6858257
 _Iyear_1992 |  -.1457678   .3858618    -0.38   0.710    -.9506614    .6591257
 _Iyear_1993 |  -.1102462   .4011624    -0.27   0.786    -.9470563    .7265639
 _Iyear_1994 |  -.0178393   .4199399    -0.04   0.967    -.8938185    .8581399
 _Iyear_1995 |  -.0250906   .3587997    -0.07   0.945    -.7735336    .7233525
 _Iyear_1996 |   -.082368   .3158888    -0.26   0.797    -.7413004    .5765644
 _Iyear_1997 |   -.145846   .3068412    -0.48   0.640    -.7859055    .4942136
 _Iyear_1998 |  -.1083473   .2735719    -0.40   0.696    -.6790082    .4623137
 _Iyear_1999 |    .014184   .2609496     0.05   0.957    -.5301473    .5585152
 _Iyear_2000 |  -.0273199   .2227771    -0.12   0.904    -.4920249     .437385
 _Iyear_2001 |    .020526   .2171885     0.09   0.926    -.4325214    .4735734
 _Iyear_2002 |    .134062   .2251193     0.60   0.558    -.3355286    .6036527
 _Iyear_2003 |   .1047764   .2058809     0.51   0.616    -.3246836    .5342364
 _Iyear_2004 |   .1346374   .1681764     0.80   0.433    -.2161726    .4854473
 _Iyear_2005 |   .0525024   .1219116     0.43   0.671    -.2018008    .3068056
 _Iyear_2006 |   .0201419   .1043943     0.19   0.849    -.1976208    .2379046
 _Iyear_2007 |   .0438741   .0598369     0.73   0.472    -.0809434    .1686916
 _Iyear_2008 |  (omitted)
 _Iyear_2009 |  -.0808973   .0484992    -1.67   0.111     -.182065    .0202703
 _Iyear_2010 |   .1161364   .0956408     1.21   0.239    -.0833669    .3156397
       _cons |  -15.23896   8.563827    -1.78   0.090    -33.10279    2.624873
-------------+----------------------------------------------------------------
     sigma_u |  1.6049458
     sigma_e |  .35420241
         rho |  .95355612   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

```
//Mean Group
All coefficients represent averages across groups (group variable: id)
Coefficient averages computed as outlier-robust means (using rreg)
Mean Group type estimation                      Number of obs      =       588
Group variable: id                              Number of groups   =        21
                                                Obs per group: min =        28
                                                               avg =      28.0
                                                               max =        28
                                                Wald chi2(8)       =     21.25
                                                Prob > chi2        =    0.0065
------------------------------------------------------------------------------
       lnFDI |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     lnGDPL1 |   .9337245   .3463503     2.70   0.007     .2548904    1.612559
     lnTAXL1 |  -.2809236    .121019    -2.32   0.020    -.5181164   -.0437308
    lnCOSTL1 |  -.0074236   .1617816    -0.05   0.963    -.3245096    .3096624
   lnSKILLL1 |   .1070559   .3414711     0.31   0.754    -.5622152     .776327
    lnOPENL1 |  -.0648992   .2519963    -0.26   0.797    -.5588029    .4290044
     lnFERL1 |  -.0391874   .0144887    -2.70   0.007    -.0675846   -.0107901
      lnTCL1 |   .0619102   .1347778     0.46   0.646    -.2022493    .3260698
    lnRISKL1 |   .3112536   .3275403     0.95   0.342    -.3307136    .9532208
       trend |    .043683    .012876     3.39   0.001     .0184465    .0689196
       _cons |  -.1945108   2.698804    -0.07   0.943    -5.484069    5.095047
------------------------------------------------------------------------------
Root Mean Squared Error (sigma): 0.1325

//CCEP
i.id              _Iid_1-21           (naturally coded; _Iid_1 omitted)
i.id|lnFDIT       _IidXlnFD_#         (coded as above)
i.id|lnGDPTL1     _IidXlnGD_#         (coded as above)
i.id|lnTAXTL1     _IidXlnTA_#         (coded as above)
i.id|lnCOSTTL1    _IidXlnCO_#         (coded as above)
i.id|lnSKILLTL1   _IidXlnSK_#         (coded as above)
i.id|lnOPENTL1    _IidXlnOP_#         (coded as above)
i.id|lnFERTL1     _IidXlnFE_#         (coded as above)
i.id|lnTCTL1      _IidXlnTC_#         (coded as above)
i.id|lnRISKTL1    _IidXlnRI_#         (coded as above)
note: lnTCTL1 omitted because of collinearity
Fixed-effects (within) regression               Number of obs      =       588
Group variable: id                              Number of groups   =        21
R-sq:  within  = 0.9732                         Obs per group: min =        28
       between = 0.1073                                        avg =      28.0
       overall = 0.0899                                        max =        28
                                                F(196,371)         =     68.62
corr(u_i, Xb)  = -0.9986                        Prob > F           =    0.0000
------------------------------------------------------------------------------
       lnFDI |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     lnGDPL1 |   .3611181   .4271416     0.85   0.398    -.4788041     1.20104
     lnTAXL1 |  -.1697191   .1225232    -1.39   0.167    -.4106462     .071208
    lnCOSTL1 |   .3971959   .2381851     1.67   0.096    -.0711663    .8655581
   lnSKILLL1 |    .002144   .3135117     0.01   0.995    -.6143388    .6186268
    lnOPENL1 |    .111506   .2628776     0.42   0.672    -.4054109    .6284229
     lnFERL1 |  -.0221923   .0116643    -1.90   0.058    -.0451288    .0007442
      lnTCL1 |   -.722752   .8524638    -0.85   0.397    -2.399019    .9535147
    lnRISKL1 |  -.3913268   .2591687    -1.51   0.132    -.9009507    .1182971
      lnFDIT |   .8647692   1.093898     0.79   0.430    -1.286248    3.015786
 _IidXlnFD_2 |   1.074117    1.56836     0.68   0.494    -2.009874    4.158108
 _IidXlnFD_3 |   .7361232   1.533177     0.48   0.631    -2.278683     3.75093
 _IidXlnFD_4 |  -.0454258   1.538482    -0.03   0.976    -3.070664    2.979812
 _IidXlnFD_5 |   .0683391   1.548169     0.04   0.965    -2.975947    3.112625
 _IidXlnFD_6 |   1.933772   1.595462     1.21   0.226     -1.20351    5.071054
 _IidXlnFD_7 |  -.3855854   1.561847    -0.25   0.805    -3.456767    2.685597
```

```
  _IidXlnFD_8 |  -.0937484   1.539119    -0.06   0.951    -3.120238    2.932742
  _IidXlnFD_9 |  -.8187847   1.681221    -0.49   0.627    -4.124701    2.487132
 _IidXlnFD_10 |  -1.727446   1.544902    -1.12   0.264    -4.765308    1.310416
 _IidXlnFD_11 |  -.0751844   1.536378    -0.05   0.961    -3.096285    2.945916
 _IidXlnFD_12 |  -.0150126   1.542758    -0.01   0.992    -3.048658    3.018633
 _IidXlnFD_13 |   1.787026   1.537855     1.16   0.246    -1.236979    4.811031
 _IidXlnFD_14 |   1.441059   1.530536     0.94   0.347    -1.568556    4.450673
 _IidXlnFD_15 |  -.8373703   1.541376    -0.54   0.587    -3.868299    2.193559
 _IidXlnFD_16 |    .130952   1.538232     0.09   0.932    -2.893794    3.155698
 _IidXlnFD_17 |   -.337674   1.558696    -0.22   0.829     -3.40266    2.727312
 _IidXlnFD_18 |   .7530994   1.528618     0.49   0.623    -2.252744    3.758942
 _IidXlnFD_19 |   .9752705   1.534186     0.64   0.525     -2.04152    3.992061
 _IidXlnFD_20 |  -.7806193   1.530549    -0.51   0.610    -3.790259     2.22902
 _IidXlnFD_21 |  -.9387065   1.527556    -0.61   0.539     -3.94246    2.065047
     lnGDPTL1 |  -1.495047   2.697381    -0.55   0.580    -6.799121    3.809026
  _IidXlnGD_2 |  -.9230403    3.86894    -0.24   0.812    -8.530842    6.684761
  _IidXlnGD_3 |  -1.850668    3.92809    -0.47   0.638     -9.57478    5.873444
  _IidXlnGD_4 |   1.275685   3.969191     0.32   0.748    -6.529247    9.080617
  _IidXlnGD_5 |   .5217547    3.89538     0.13   0.894    -7.138037    8.181547
  _IidXlnGD_6 |  -6.020173    4.04401    -1.49   0.137    -13.97223    1.931882
  _IidXlnGD_7 |   2.879254   3.816153     0.75   0.451    -4.624748    10.38326
  _IidXlnGD_8 |   1.581385   3.863206     0.41   0.683    -6.015141    9.177911
  _IidXlnGD_9 |   1.037854    4.29145     0.24   0.809    -7.400763    9.476471
 _IidXlnGD_10 |   7.863764    4.01329     1.96   0.051    -.0278839    15.75541
 _IidXlnGD_11 |   1.828489   3.835006     0.48   0.634    -5.712587    9.369564
 _IidXlnGD_12 |   2.535835   3.889826     0.65   0.515    -5.113035    10.18471
 _IidXlnGD_13 |  -.9033872   3.882035    -0.23   0.816    -8.536938    6.730164
 _IidXlnGD_14 |  -1.970146   3.846495    -0.51   0.609    -9.533813    5.593521
 _IidXlnGD_15 |   2.201788   3.837628     0.57   0.566    -5.344443     9.74802
 _IidXlnGD_16 |  -1.383831   3.837262    -0.36   0.719    -8.929342     6.16168
 _IidXlnGD_17 |   4.181273   3.912861     1.07   0.286    -3.512894    11.87544
 _IidXlnGD_18 |   1.922777   3.852748     0.50   0.618    -5.653185    9.498739
 _IidXlnGD_19 |    .463798   3.831114     0.12   0.904    -7.069624     7.99722
 _IidXlnGD_20 |   4.645019   3.816453     1.22   0.224    -2.859574    12.14961
 _IidXlnGD_21 |   3.292594   3.828219     0.86   0.390    -4.235134    10.82032
     lnTAXTL1 |  -.6483291   .9519916    -0.68   0.496    -2.520305    1.223647
  _IidXlnTA_2 |   .0492587   1.345497     0.04   0.971    -2.596497    2.695015
  _IidXlnTA_3 |   .1043987   1.437044     0.07   0.942    -2.721374    2.930171
  _IidXlnTA_4 |   .9550086   1.333959     0.72   0.474     -1.66806    3.578078
  _IidXlnTA_5 |   .7295006   1.361646     0.54   0.592     -1.94801    3.407012
  _IidXlnTA_6 |    2.35819   1.388454     1.70   0.090    -.3720367    5.088416
  _IidXlnTA_7 |   .9973145   1.332219     0.75   0.455    -1.622332    3.616961
  _IidXlnTA_8 |   1.050908   1.359333     0.77   0.440    -1.622055    3.723871
  _IidXlnTA_9 |  -1.209409   1.489742    -0.81   0.417    -4.138805    1.719988
 _IidXlnTA_10 |  -.3152447   1.370489    -0.23   0.818    -3.010145    2.379656
 _IidXlnTA_11 |   2.450563   1.348319     1.82   0.070    -.2007418    5.101868
 _IidXlnTA_12 |   1.150643   1.420845     0.81   0.419    -1.643278    3.944563
 _IidXlnTA_13 |  -.5358673   1.360757    -0.39   0.694    -3.211631    2.139897
 _IidXlnTA_14 |   1.544755   1.328607     1.16   0.246    -1.067791      4.1573
 _IidXlnTA_15 |   1.059558   1.352638     0.78   0.434    -1.600242    3.719357
 _IidXlnTA_16 |  -.2299846   1.359234    -0.17   0.866    -2.902754    2.442784
 _IidXlnTA_17 |   3.031786   1.387379     2.19   0.029     .3036727    5.759899
 _IidXlnTA_18 |   2.112824   1.338241     1.58   0.115    -.5186639    4.744313
 _IidXlnTA_19 |   1.206638   1.382541     0.87   0.383     -1.51196    3.925237
 _IidXlnTA_20 |   .1457851   1.340603     0.11   0.913    -2.490348    2.781919
 _IidXlnTA_21 |   .2778047   1.372517     0.20   0.840    -2.421085    2.976694
    lnCOSTTL1 |   .5399722   .6478116     0.83   0.405    -.7338708    1.813815
  _IidXlnCO_2 |  -.7724987   .7762691    -1.00   0.320    -2.298938    .7539403
  _IidXlnCO_3 |  -.6888745   .7506082    -0.92   0.359    -2.164855    .7871056
  _IidXlnCO_4 |  -1.059598   .7791694    -1.36   0.175    -2.591741    .4725437
  _IidXlnCO_5 |  -1.098016   .7354975    -1.49   0.136    -2.544283    .3482506
  _IidXlnCO_6 |  -1.481916   .7945764    -1.87   0.063    -3.044355     .080522
  _IidXlnCO_7 |  -.4707906   .7736184    -0.61   0.543    -1.992017    1.050436
```

```
 _IidXlnCO_8 |  -1.062325   .7539381    -1.41   0.160    -2.544853    .4202034
 _IidXlnCO_9 |   .0461204   .7746982     0.06   0.953     -1.47723    1.569471
_IidXlnCO_10 |  -1.504738   .7933295    -1.90   0.059    -3.064725    .0552481
_IidXlnCO_11 |  -1.548977   .7357196    -2.11   0.036     -2.99568   -.1022732
_IidXlnCO_12 |  -.9965192   .8011001    -1.24   0.214    -2.571785     .578747
_IidXlnCO_13 |  -1.231598   .7375983    -1.67   0.096    -2.681996    .2187998
_IidXlnCO_14 |  -.6115377   .7654382    -0.80   0.425    -2.116679    .8936037
_IidXlnCO_15 |   .3628905   .7755552     0.47   0.640    -1.162145    1.887926
_IidXlnCO_16 |   .5446965   .7401934     0.74   0.462    -.9108042    2.000197
_IidXlnCO_17 |  -1.569051   .7611218    -2.06   0.040    -3.065705   -.0723972
_IidXlnCO_18 |  -1.790185   .8000277    -2.24   0.026    -3.363342    -.217027
_IidXlnCO_19 |  -2.616833   .7317102    -3.58   0.000    -4.055652   -1.178013
_IidXlnCO_20 |  -1.743917   .7415813    -2.35   0.019    -3.202147   -.2856874
_IidXlnCO_21 |  -.3618078   .7463644    -0.48   0.628    -1.829443    1.105827
  lnSKILLTL1 |   .4589098   1.649642     0.28   0.781     -2.78491     3.70273
 _IidXlnSK_2 |  -.8962645   2.232565    -0.40   0.688    -5.286333    3.493804
 _IidXlnSK_3 |  -.6976459   2.442115    -0.29   0.775    -5.499769    4.104477
 _IidXlnSK_4 |  -.4587543   2.287577    -0.20   0.841    -4.956997    4.039488
 _IidXlnSK_5 |   -1.89631   2.236637    -0.85   0.397    -6.294386    2.501767
 _IidXlnSK_6 |   .5192336   2.312979     0.22   0.823    -4.028959    5.067426
 _IidXlnSK_7 |  -.8682793   2.651861    -0.33   0.744    -6.082843    4.346284
 _IidXlnSK_8 |  -2.795725   2.330001    -1.20   0.231     -7.37739    1.785941
 _IidXlnSK_9 |   3.389335     2.5588     1.32   0.186    -1.642236    8.420905
_IidXlnSK_10 |  -1.543376   2.289425    -0.67   0.501    -6.045254    2.958502
_IidXlnSK_11 |  -1.335397   2.226497    -0.60   0.549    -5.713534    3.042739
_IidXlnSK_12 |  -2.080725   2.270784    -0.92   0.360    -6.545946    2.384496
_IidXlnSK_13 |   .8941616   2.252803     0.40   0.692    -3.535703    5.324026
_IidXlnSK_14 |  -1.457289   2.192115    -0.66   0.507    -5.767817    2.853238
_IidXlnSK_15 |   5.065931   2.202242     2.30   0.022     .7354896    9.396372
_IidXlnSK_16 |  -2.665521   2.220953    -1.20   0.231    -7.032755    1.701713
_IidXlnSK_17 |    -.48511   2.313764    -0.21   0.834    -5.034847    4.064627
_IidXlnSK_18 |  -1.799651   2.179605    -0.83   0.410    -6.085581    2.486279
_IidXlnSK_19 |  -.7299224   2.337002    -0.31   0.755    -5.325353    3.865508
_IidXlnSK_20 |  -1.478153   2.264388    -0.65   0.514    -5.930798    2.974491
_IidXlnSK_21 |   1.141152   2.309376     0.49   0.622    -3.399956    5.682261
   lnOPENTL1 |   .2092925   1.656578     0.13   0.900    -3.048167    3.466752
 _IidXlnOP_2 |  -.6282647   2.063535    -0.30   0.761    -4.685955    3.429426
 _IidXlnOP_3 |   .4226466   2.248769     0.19   0.851    -3.999285    4.844579
 _IidXlnOP_4 |  -1.264846   2.019894    -0.63   0.532    -5.236724    2.707031
 _IidXlnOP_5 |   .7039365   2.024563     0.35   0.728    -3.277122    4.684995
 _IidXlnOP_6 |   1.248455   2.106648     0.59   0.554    -2.894012    5.390922
 _IidXlnOP_7 |  -.3165895   2.034306    -0.16   0.876    -4.316806    3.683627
 _IidXlnOP_8 |  -.4078095   2.032753    -0.20   0.841    -4.404972    3.589353
 _IidXlnOP_9 |   .1524799   2.098515     0.07   0.942    -3.973995    4.278955
_IidXlnOP_10 |  -.1668731   2.023799    -0.08   0.934    -4.146428    3.812682
_IidXlnOP_11 |  -1.198781   2.036544    -0.59   0.556    -5.203397    2.805835
_IidXlnOP_12 |  -1.478381   2.178814    -0.68   0.498    -5.762756    2.805993
_IidXlnOP_13 |    .257604   2.080654     0.12   0.902     -3.83375    4.348958
_IidXlnOP_14 |     1.7343   2.015027     0.86   0.390    -2.228005    5.696606
_IidXlnOP_15 |  -.2661102   2.035714    -0.13   0.896    -4.269094    3.736874
_IidXlnOP_16 |   .8022899   2.028598     0.40   0.693    -3.186701    4.791281
_IidXlnOP_17 |  -1.347482   2.080769    -0.65   0.518    -5.439063    2.744098
_IidXlnOP_18 |  -2.261378   2.072815    -1.09   0.276    -6.337317    1.814561
_IidXlnOP_19 |   1.349758   2.064271     0.65   0.514     -2.70938    5.408896
_IidXlnOP_20 |  -2.631257   2.063315    -1.28   0.203    -6.688516    1.426002
_IidXlnOP_21 |   .1212661   2.044023     0.06   0.953    -3.898057    4.140589
     lnFERTL1 |  -.0260971   .1319125    -0.20   0.843    -.2854871    .2332928
 _IidXlnFE_2 |    .077937   .1823683     0.43   0.669    -.2806682    .4365422
 _IidXlnFE_3 |   .0350623   .1880066     0.19   0.852    -.3346299    .4047545
 _IidXlnFE_4 |   .0760678   .1772959     0.43   0.668     -.272563    .4246986
 _IidXlnFE_5 |   .0964402   .1784497     0.54   0.589    -.2544596    .4473399
 _IidXlnFE_6 |  -.0166957   .1781378    -0.09   0.925    -.3669821    .3335907
 _IidXlnFE_7 |   .0477103   .1796726     0.27   0.791    -.3055941    .4010148
```

```
 _IidXlnFE_8 |   .0922653   .1782938     0.52   0.605     -.258328    .4428585
 _IidXlnFE_9 |  -.1368933   .1955719    -0.70   0.484    -.5214618    .2476752
_IidXlnFE_10 |  -.1174079   .1782246    -0.66   0.510    -.4678649    .2330491
_IidXlnFE_11 |   .0974604   .1780271     0.55   0.584    -.2526082     .447529
_IidXlnFE_12 |   .0448684    .180918     0.25   0.804    -.3108848    .4006217
_IidXlnFE_13 |   .1715298   .1821657     0.94   0.347    -.1866769    .5297366
_IidXlnFE_14 |   .0574895   .1778232     0.32   0.747    -.2921782    .4071572
_IidXlnFE_15 |    .062701   .1774945     0.35   0.724    -.2863204    .4117223
_IidXlnFE_16 |   .0900766   .1785721     0.50   0.614    -.2610637    .4412169
_IidXlnFE_17 |   .0473052   .1789223     0.26   0.792    -.3045238    .3991341
_IidXlnFE_18 |   .0734447   .1782495     0.41   0.681    -.2770613    .4239508
_IidXlnFE_19 |  -.0215163   .1801938    -0.12   0.905    -.3758456    .3328131
_IidXlnFE_20 |   .0925627   .1789165     0.52   0.605    -.2592549    .4443803
_IidXlnFE_21 |  -.0363362    .180013    -0.20   0.840    -.3903101    .3176376
     lnTCTL1 |  (omitted)
 _IidXlnTC_2 |   .9800356   .9509938     1.03   0.303    -.8899784     2.85005
 _IidXlnTC_3 |   .0312249   1.001294     0.03   0.975    -1.937698    2.000147
 _IidXlnTC_4 |   1.488001   1.427738     1.04   0.298    -1.319473    4.295475
 _IidXlnTC_5 |    .631366   .8908404     0.71   0.479    -1.120364    2.383096
 _IidXlnTC_6 |    .361675   .9363491     0.39   0.700    -1.479542    2.202892
 _IidXlnTC_7 |   1.334778   1.271731     1.05   0.295    -1.165927    3.835482
 _IidXlnTC_8 |   1.027244   .9184169     1.12   0.264    -.7787111      2.8332
 _IidXlnTC_9 |   .6456263   .8747871     0.74   0.461    -1.074536    2.365789
_IidXlnTC_10 |   1.350612   1.016112     1.33   0.185    -.6474496    3.348673
_IidXlnTC_11 |  -.0566186   1.334385    -0.04   0.966    -2.680525    2.567287
_IidXlnTC_12 |   .9387007   2.133639     0.44   0.660    -3.256841    5.134242
_IidXlnTC_13 |   .5382173   .8727412     0.62   0.538    -1.177922    2.254357
_IidXlnTC_14 |   1.480101   1.118793     1.32   0.187    -.7198687    3.680071
_IidXlnTC_15 |   .9971205   .9980376     1.00   0.318    -.9653995     2.95964
_IidXlnTC_16 |   .1807277   .9592295     0.19   0.851    -1.705481    2.066936
_IidXlnTC_17 |   .6729606   .8752998     0.77   0.442     -1.04821    2.394132
_IidXlnTC_18 |   .5377023   1.030822     0.52   0.602    -1.489284    2.564689
_IidXlnTC_19 |   .8762771     .94638     0.93   0.355    -.9846645    2.737219
_IidXlnTC_20 |   .9004704   .9027221     1.00   0.319    -.8746232    2.675564
_IidXlnTC_21 |   .6640898   1.068911     0.62   0.535    -1.437795    2.765974
   lnRISKTL1 |   2.168175   2.351145     0.92   0.357    -2.455067    6.791416
 _IidXlnRI_2 |  -.5427674   3.346953    -0.16   0.871    -7.124146    6.038611
 _IidXlnRI_3 |   -1.09866   3.478024    -0.32   0.752    -7.937772    5.740453
 _IidXlnRI_4 |  -.1949755   3.326273    -0.06   0.953    -6.735689    6.345738
 _IidXlnRI_5 |  -1.424027   3.258722    -0.44   0.662    -7.831909    4.983855
 _IidXlnRI_6 |   -2.76437   3.646483    -0.76   0.449    -9.934737    4.405998
 _IidXlnRI_7 |  -4.490562   3.336757    -1.35   0.179    -11.05189    2.070766
 _IidXlnRI_8 |  -4.821546   3.371345    -1.43   0.154    -11.45089    1.807795
 _IidXlnRI_9 |   3.134747   3.527079     0.89   0.375    -3.800827    10.07032
_IidXlnRI_10 |    1.73597   3.413134     0.51   0.611    -4.975544    8.447483
_IidXlnRI_11 |  -3.897641   3.395851    -1.15   0.252    -10.57517    2.779888
_IidXlnRI_12 |  -3.713816   3.449729    -1.08   0.282    -10.49729    3.069658
_IidXlnRI_13 |  -.4446616   3.400621    -0.13   0.896     -7.13157    6.242247
_IidXlnRI_14 |  -2.362135   3.395194    -0.70   0.487    -9.038373    4.314104
_IidXlnRI_15 |  -5.768022   3.410735    -1.69   0.092    -12.47482    .9387753
_IidXlnRI_16 |  -1.205867   3.395447    -0.36   0.723    -7.882603    5.470868
_IidXlnRI_17 |  -6.358076   3.388912    -1.88   0.061    -13.02196    .3058095
_IidXlnRI_18 |  -.4255526   3.372212    -0.13   0.900    -7.056599    6.205494
_IidXlnRI_19 |  -1.575793   3.524897    -0.45   0.655    -8.507077     5.35549
_IidXlnRI_20 |   .3732328   3.306242     0.11   0.910    -6.128092    6.874558
_IidXlnRI_21 |  -.2490021   3.533098    -0.07   0.944     -7.19641    6.698406
       _cons |   .2610797   6.489765     0.04   0.968    -12.50026    13.02242
-------------+----------------------------------------------------------------
     sigma_u |  31.373863
     sigma_e |   .1692089
         rho |  .99997091   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(20, 371) =     1.01             Prob > F = 0.4478
```

```
//CCEMG
All coefficients represent averages across groups (group variable: id)
Coefficient averages computed as outlier-robust means (using rreg)
Mean Group type estimation                       Number of obs      =       588
Group variable: id                               Number of groups   =        21
                                                 Obs per group: min =        28
                                                                avg =      28.0
                                                                max =        28
                                                 Wald chi2(8)       =      5.95
                                                 Prob > chi2        =    0.6530
------------------------------------------------------------------------------
       lnFDI |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lnGDPL1 |   1.052306   .9057264     1.16   0.245    -.7228853    2.827497
      lnTAXL1 |  -.0784741   .1936484    -0.41   0.685    -.4580181    .3010699
     lnCOSTL1 |   .4110899    .493875     0.83   0.405    -.5568873    1.379067
    lnSKILLL1 |  -.2894749   .5085488    -0.57   0.569    -1.286212    .7072625
     lnOPENL1 |  -.3615503   .2795354    -1.29   0.196    -.9094295     .186329
      lnFERL1 |  -.0197208   .0250113    -0.79   0.430    -.0687421    .0293006
       lnTCL1 |    .099973   .1333873     0.75   0.454    -.1614612    .3614073
     lnRISKL1 |    .400356   .5342838     0.75   0.454    -.6468211    1.447533
    lnFDI_avg |    1.11874   .4395406     2.55   0.011     .2572561    1.980224
  lnGDPL1_avg |  -1.336814   1.236547    -1.08   0.280    -3.760401    1.086773
  lnTAXL1_avg |   .6096431   .3438605     1.77   0.076    -.0643111    1.283597
 lnCOSTL1_avg |  -.6559257   .5795141    -1.13   0.258    -1.791752     .479901
 lnSKILLL1_~g |    -.60674   .6254257    -0.97   0.332    -1.832552     .619072
 lnOPENL1_avg |  -.2090073   .7036553    -0.30   0.766    -1.588146    1.170132
  lnFERL1_avg |   .0311635    .025192     1.24   0.216    -.0182119    .0805389
   lnTCL1_avg |  -.1141706   .2770761    -0.41   0.680    -.6572298    .4288886
 lnRISKL1_avg |  -.9391758   .8781099    -1.07   0.285     -2.66024     .781888
        _cons |  -12.86569   3.747706    -3.43   0.001    -20.21106   -5.520323
------------------------------------------------------------------------------
Root Mean Squared Error (sigma): 0.0890
```

Appendix I.5:

STATA output of regressions of the models using two-year lagged values for independent variables

```
//Pooled OLS
i.year            _Iyear_1982-2010    (naturally coded; _Iyear_1982 omitted)
note: _Iyear_1983 omitted because of collinearity
note: _Iyear_2010 omitted because of collinearity
Linear regression                                      Number of obs =     567
                                                       F( 19,    20) =       .
                                                       Prob > F      =       .
                                                       R-squared     =  0.7556
                                                       Root MSE      =  .87333
                                   (Std. Err. adjusted for 21 clusters in id)
------------------------------------------------------------------------------
             |               Robust
       lnFDI |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lnGDPL2 |   1.446759   .1720259     8.41   0.000     1.087919    1.805598
      lnTAXL2 |  -.7185231   .1679924    -4.28   0.000    -1.068949   -.3680971
     lnCOSTL2 |  -1.977598   .7308324    -2.71   0.014    -3.502087   -.4531082
    lnSKILLL2 |   .9945309   1.171355     0.85   0.406    -1.448873    3.437935
     lnOPENL2 |    1.67061   .5410285     3.09   0.006     .5420446    2.799176
      lnFERL2 |  -.0663252   .0926897    -0.72   0.483    -.2596725    .1270221
       lnTCL2 |  -.5438219   .4633337    -1.17   0.254    -1.510319    .4226752
     lnRISKL2 |   1.987067   1.154749     1.72   0.101    -.4216976    4.395832
 _Iyear_1983 |  (omitted)
 _Iyear_1984 |  -.6064064   .3215852    -1.89   0.074    -1.277221    .0644085
 _Iyear_1985 |  -.7574061   .3280504    -2.31   0.032    -1.441707   -.0731049
 _Iyear_1986 |  -1.002393    .370315    -2.71   0.014    -1.774857   -.2299295
 _Iyear_1987 |  -.9193704    .390053    -2.36   0.029    -1.733007    -.105734
 _Iyear_1988 |  -.3955445   .3273194    -1.21   0.241    -1.078321    .2872318
 _Iyear_1989 |   .0095037   .3583258     0.03   0.979    -.7379508    .7569581
 _Iyear_1990 |   .1816752    .379602     0.48   0.637    -.6101607     .973511
 _Iyear_1991 |  -.0766061   .3658901    -0.21   0.836    -.8398395    .6866274
 _Iyear_1992 |   .2016471   .3273924     0.62   0.545    -.4812815    .8845758
 _Iyear_1993 |   .2235833    .358701     0.62   0.540    -.5246539    .9718205
 _Iyear_1994 |   .3306558   .3324517     0.99   0.332    -.3628263    1.024138
 _Iyear_1995 |   .1461187   .3711146     0.39   0.698    -.6280127    .9202501
 _Iyear_1996 |   .0976778   .3743799     0.26   0.797     -.683265    .8786207
 _Iyear_1997 |   .1883087   .3287021     0.57   0.573    -.4973518    .8739693
 _Iyear_1998 |   .2062839    .315377     0.65   0.521     -.451581    .8641488
 _Iyear_1999 |  -.0633298   .3630044    -0.17   0.863    -.8205438    .6938842
 _Iyear_2000 |  -.1787825   .3797126    -0.47   0.643    -.9708491    .6132841
 _Iyear_2001 |  -.2653926   .4172737    -0.64   0.532     -1.13581     .605025
 _Iyear_2002 |  -.6230466   .3931773    -1.58   0.129      -1.4432    .1971068
 _Iyear_2003 |  -.7897169   .3886244    -2.03   0.056    -1.600373    .0209394
 _Iyear_2004 |  -.4966694   .3877756    -1.28   0.215    -1.305555    .3122164
 _Iyear_2005 |  -.1661217   .3386248    -0.49   0.629    -.8724807    .5402373
 _Iyear_2006 |   .0294854   .2166469     0.14   0.893    -.4224322    .4814029
 _Iyear_2007 |  -.0047519   .1751904    -0.03   0.979    -.3701927    .3606889
 _Iyear_2008 |  -.1363905   .1401095    -0.97   0.342    -.4286539    .1558728
 _Iyear_2009 |  -.0373946   .0504196    -0.74   0.467    -.1425681    .0677788
 _Iyear_2010 |  (omitted)
       _cons |  -13.89839   4.027941    -3.45   0.003    -22.30053   -5.496254
------------------------------------------------------------------------------
```

```
//Fixed Effects
i.year           _Iyear_1982-2010    (naturally coded; _Iyear_1982 omitted)
note: _Iyear_1983 omitted because of collinearity
note: _Iyear_1984 omitted because of collinearity
Fixed-effects (within) regression               Number of obs      =       567
Group variable: id                              Number of groups   =        21
R-sq:  within  = 0.8303                         Obs per group: min =        27
       between = 0.4834                                        avg =      27.0
       overall = 0.5191                                        max =        27
                                                F(20,20)           =         .
corr(u_i, Xb)  = -0.6569                        Prob > F           =         .
                                    (Std. Err. adjusted for 21 clusters in id)
------------------------------------------------------------------------------
             |               Robust
       lnFDI |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     lnGDPL2 |    1.97435   .7533876     2.62   0.016     .4028114    3.545889
     lnTAXL2 |  -.1735896   .2376459    -0.73   0.474    -.6693103    .3221311
    lnCOSTL2 |   .6582673   .4414235     1.49   0.152     -.262526    1.579061
   lnSKILLL2 |   .7244202   .6766258     1.07   0.297    -.6869964    2.135837
    lnOPENL2 |   1.265146   .6054251     2.09   0.050     .0022518    2.528041
     lnFERL2 |    -.04861   .0202692    -2.40   0.026    -.0908908   -.0063293
      lnTCL2 |   .4219347   .2857149     1.48   0.155    -.1740562    1.017926
    lnRISKL2 |   .2992198   .4099917     0.73   0.474     -.556008    1.154448
 _Iyear_1983 |  (omitted)
 _Iyear_1984 |  (omitted)
 _Iyear_1985 |   .1222867   .0499308     2.45   0.024      .018133    .2264405
 _Iyear_1986 |   .1506381   .0764997     1.97   0.063    -.0089375    .3102137
 _Iyear_1987 |   .2938978   .1021483     2.88   0.009     .0808201    .5069755
 _Iyear_1988 |   .2087684   .1126323     1.85   0.079    -.0261783    .4437152
 _Iyear_1989 |   .0866774   .1659495     0.52   0.607    -.2594871    .4328419
 _Iyear_1990 |   .2012451   .2254505     0.89   0.383    -.2690365    .6715266
 _Iyear_1991 |   .1164839   .2465294     0.47   0.642    -.3977674    .6307352
 _Iyear_1992 |   .0274742   .2742274     0.10   0.921    -.5445541    .5995026
 _Iyear_1993 |   .0839832   .2562592     0.33   0.747    -.4505642    .6185306
 _Iyear_1994 |   .1512573   .2651325     0.57   0.575    -.4017995     .704314
 _Iyear_1995 |   .2936794   .2893461     1.01   0.322     -.309886    .8972447
 _Iyear_1996 |   .1743779   .3328717     0.52   0.606    -.5199804    .8687361
 _Iyear_1997 |   .0777809   .3740812     0.21   0.837    -.7025389    .8581007
 _Iyear_1998 |   .1455425   .3974722     0.37   0.718      -.68357     .974655
 _Iyear_1999 |    .171086   .4163307     0.41   0.685    -.6973646    1.039537
 _Iyear_2000 |   .1995212   .4221668     0.47   0.642    -.6811032    1.080146
 _Iyear_2001 |   .1701631   .4511391     0.38   0.710    -.7708966    1.111223
 _Iyear_2002 |   .2432041   .4656433     0.52   0.607    -.7281108    1.214519
 _Iyear_2003 |   .2978997   .4711708     0.63   0.534    -.6849454    1.280745
 _Iyear_2004 |   .3529805    .492639     0.72   0.482    -.6746465    1.380608
 _Iyear_2005 |   .2411886   .5140983     0.47   0.644    -.8312016    1.313579
 _Iyear_2006 |   .2172267   .5361818     0.41   0.690     -.901229    1.335682
 _Iyear_2007 |   .2321796   .5605504     0.41   0.683    -.9371081    1.401467
 _Iyear_2008 |   .2066069   .5983067     0.35   0.733    -1.041439    1.454653
 _Iyear_2009 |   .1272656   .6154083     0.21   0.838    -1.156454    1.410985
 _Iyear_2010 |   .0165833   .6339046     0.03   0.979    -1.305718    1.338885
       _cons |  -15.68424   8.936365    -1.76   0.095    -34.32517    2.956688
-------------+----------------------------------------------------------------
     sigma_u |  1.5481811
     sigma_e |  .34897698
         rho |  .95164673   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

```
//Mean Group
All coefficients represent averages across groups (group variable: id)
Coefficient averages computed as outlier-robust means (using rreg)
Mean Group type estimation                      Number of obs      =       567
Group variable: id                              Number of groups   =        21
                                                Obs per group: min =        27
                                                               avg =      27.0
                                                               max =        27
                                                Wald chi2(8)       =     19.55
                                                Prob > chi2        =    0.0122
------------------------------------------------------------------------------
       lnFDI |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     lnGDPL2 |    1.01813    .605174     1.68   0.092     -.167989    2.204249
     lnTAXL2 |  -.0754065   .1458581    -0.52   0.605    -.3612831    .2104701
    lnCOSTL2 |  -.4021292   .1832275    -2.19   0.028    -.7612485    -.04301
   lnSKILLL2 |   .3230671   .4594356     0.70   0.482    -.5774101    1.223544
    lnOPENL2 |   -.314211   .3325999    -0.94   0.345    -.9660949    .3376728
     lnFERL2 |  -.0413903   .0159554    -2.59   0.009    -.0726624   -.0101183
      lnTCL2 |   .0698565    .145751     0.48   0.632    -.2158102    .3555233
    lnRISKL2 |   .4505131   .2483912     1.81   0.070    -.0363247    .9373509
       trend |   .0615888   .0169257     3.64   0.000     .0284151    .0947625
       _cons |  -3.428185   6.242926    -0.55   0.583    -15.66409    8.807724
------------------------------------------------------------------------------
Root Mean Squared Error (sigma): 0.1236

//CCEP
i.id              _Iid_1-21           (naturally coded; _Iid_1 omitted)
i.id|lnFDIT       _IidXlnFD_#         (coded as above)
i.id|lnGDPTL2     _IidXlnGD_#         (coded as above)
i.id|lnTAXTL2     _IidXlnTA_#         (coded as above)
i.id|lnCOSTTL2    _IidXlnCO_#         (coded as above)
i.id|lnSKILLTL2   _IidXlnSK_#         (coded as above)
i.id|lnOPENTL2    _IidXlnOP_#         (coded as above)
i.id|lnFERTL2     _IidXlnFE_#         (coded as above)
i.id|lnTCTL2      _IidXlnTC_#         (coded as above)
i.id|lnRISKTL2    _IidXlnRI_#         (coded as above)
Fixed-effects (within) regression               Number of obs      =       567
Group variable: id                              Number of groups   =        21
R-sq:  within  = 0.9831                         Obs per group: min =        27
       between = 0.0670                                        avg =      27.0
       overall = 0.0572                                        max =        27
                                                F(197,349)         =    103.36
corr(u_i, Xb)  = -0.9988                        Prob > F           =    0.0000
------------------------------------------------------------------------------
       lnFDI |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     lnGDPL2 |    .769508   .3827882     2.01   0.045     .0166462     1.52237
     lnTAXL2 |  -.3359052   .1076128    -3.12   0.002    -.5475563   -.1242541
    lnCOSTL2 |    .111441   .2117662     0.53   0.599    -.3050574    .5279394
   lnSKILLL2 |   .5828283   .2742397     2.13   0.034     .0434579    1.122199
    lnOPENL2 |   .4046675   .2461658     1.64   0.101    -.0794877    .8888227
     lnFERL2 |  -.0274833   .0095102    -2.89   0.004    -.0461878   -.0087789
      lnTCL2 |   .0341075   .0763654     0.45   0.655    -.1160867    .1843018
    lnRISKL2 |   .2659402   .2339504     1.14   0.256    -.1941899    .7260703
      lnFDIT |    .803759   .5716874     1.41   0.161    -.3206269    1.928145
 _IidXlnFD_2 |   1.738726   .8127263     2.14   0.033     .1402682    3.337183
 _IidXlnFD_3 |   1.031468   .8234758     1.25   0.211    -.5881314    2.651068
 _IidXlnFD_4 |  -1.053343   .7949126    -1.33   0.186    -2.616765    .5100789
 _IidXlnFD_5 |  -.7724526   .7994933    -0.97   0.335    -2.344884    .7999785
 _IidXlnFD_6 |   2.655309   .7948771     3.34   0.001     1.091957    4.218661
 _IidXlnFD_7 |   .5983959   .8034069     0.74   0.457    -.9817324    2.178524
 _IidXlnFD_8 |   .2297786   .8152762     0.28   0.778    -1.373694    1.833251
```

```
  _IidXlnFD_9 |  -.7157311   .8274783    -0.86   0.388    -2.343203    .9117403
 _IidXlnFD_10 |  -.4896485   .8142556    -0.60   0.548    -2.091114    1.111817
 _IidXlnFD_11 |   .4927974   .7980638     0.62   0.537    -1.076822    2.062417
 _IidXlnFD_12 |   .3241035   .8030961     0.40   0.687    -1.255413     1.90362
 _IidXlnFD_13 |  -.5391123   .7920393    -0.68   0.497    -2.096883    1.018658
 _IidXlnFD_14 |   1.664173   .8071544     2.06   0.040     .0766745    3.251672
 _IidXlnFD_15 |    .843905   .8133324     1.04   0.300    -.7557445    2.443555
 _IidXlnFD_16 |   .2849263   .8179352     0.35   0.728    -1.323776    1.893629
 _IidXlnFD_17 |  -.1043612   .8376088    -0.12   0.901    -1.751757    1.543035
 _IidXlnFD_18 |  -.2547224   .8132338    -0.31   0.754    -1.854178    1.344733
 _IidXlnFD_19 |  -.2648693   .7933988    -0.33   0.739    -1.825314    1.295575
 _IidXlnFD_20 |  -1.093056   .8010892    -1.36   0.173    -2.668626    .4825139
 _IidXlnFD_21 |  -.4552581   .7954932    -0.57   0.567    -2.019822    1.109306
     lnGDPTL2 |  -1.091033   1.856279    -0.59   0.557    -4.741933    2.559867
  _IidXlnGD_2 |  -2.934775   2.655026    -1.11   0.270     -8.15664    2.287089
  _IidXlnGD_3 |  -1.558352   2.720792    -0.57   0.567    -6.909563     3.79286
  _IidXlnGD_4 |   2.235064   2.549149     0.88   0.381    -2.778563    7.248691
  _IidXlnGD_5 |   .5274651    2.61867     0.20   0.840    -4.622894    5.677825
  _IidXlnGD_6 |  -7.968091   2.600363    -3.06   0.002    -13.08245   -2.853737
  _IidXlnGD_7 |  -.6633807   2.560684    -0.26   0.796    -5.699694    4.372933
  _IidXlnGD_8 |  -.5505827   2.674043    -0.21   0.837     -5.80985    4.708685
  _IidXlnGD_9 |   1.817338   2.755139     0.66   0.510    -3.601426    7.236103
 _IidXlnGD_10 |   1.568255   2.686218     0.58   0.560    -3.714956    6.851466
 _IidXlnGD_11 |   .1075906   2.553965     0.04   0.966    -4.915509     5.13069
 _IidXlnGD_12 |   .4779906    2.57102     0.19   0.853    -4.578651    5.534633
 _IidXlnGD_13 |   4.917629   2.620397     1.88   0.061    -.2361283    10.07139
 _IidXlnGD_14 |  -3.323165   2.576556    -1.29   0.198    -8.390696    1.744365
 _IidXlnGD_15 |  -2.488519   2.609765    -0.95   0.341    -7.621365    2.644328
 _IidXlnGD_16 |  -2.707675   2.662224    -1.02   0.310    -7.943696    2.528346
 _IidXlnGD_17 |     2.6996   2.696126     1.00   0.317    -2.603099    8.002298
 _IidXlnGD_18 |   4.884585   2.610378     1.87   0.062    -.2494669    10.01864
 _IidXlnGD_19 |   3.077586   2.595517     1.19   0.237    -2.027237    8.182408
 _IidXlnGD_20 |   4.535132   2.595249     1.75   0.081    -.5691632    9.639428
 _IidXlnGD_21 |   2.098459   2.577805     0.81   0.416    -2.971528    7.168446
     lnTAXTL2 |  -.7881775   .5703777    -1.38   0.168    -1.909988    .3336325
  _IidXlnTA_2 |   .1256738   .7887389     0.16   0.873    -1.425606    1.676953
  _IidXlnTA_3 |   .5343537     .81053     0.66   0.510    -1.059784    2.128492
  _IidXlnTA_4 |    .245523   .7904589     0.31   0.756    -1.309139    1.800185
  _IidXlnTA_5 |   .4900586   .7997481     0.61   0.540    -1.082874    2.062991
  _IidXlnTA_6 |   1.619683    .813059     1.99   0.047     .0205716    3.218795
  _IidXlnTA_7 |   2.031485   .7871798     2.58   0.010     .4832723    3.579699
  _IidXlnTA_8 |    1.43655   .8053504     1.78   0.075    -.1474006    3.020501
  _IidXlnTA_9 |  -.5801097   .8176104    -0.71   0.478    -2.188173    1.027954
 _IidXlnTA_10 |   1.509799   .8092866     1.87   0.063    -.0818933    3.101492
 _IidXlnTA_11 |   2.547232    .812526     3.13   0.002      .949168    4.145295
 _IidXlnTA_12 |   1.839628   .8054524     2.28   0.023     .2554766    3.423779
 _IidXlnTA_13 |  -.3159202   .8103519    -0.39   0.697    -1.909708    1.277867
 _IidXlnTA_14 |   .8941445   .7893973     1.13   0.258    -.6584299    2.446719
 _IidXlnTA_15 |   2.883219   .8106177     3.56   0.000     1.288908    4.477529
 _IidXlnTA_16 |   .3160387   .8016383     0.39   0.694    -1.260611    1.892689
 _IidXlnTA_17 |    3.52611   .8399738     4.20   0.000     1.874062    5.178157
 _IidXlnTA_18 |    2.11301   .7898575     2.68   0.008     .5595309     3.66649
 _IidXlnTA_19 |   .2394833   .8138611     0.29   0.769    -1.361206    1.840173
 _IidXlnTA_20 |   .8305762   .7900355     1.05   0.294    -.7232535    2.384406
 _IidXlnTA_21 |   1.319213   .8039389     1.64   0.102    -.2619619    2.900387
    lnCOSTTL2 |   .5611746   .4299008     1.31   0.193    -.2843478    1.406697
  _IidXlnCO_2 |  -.5367521   .4878255    -1.10   0.272      -1.4962    .4226957
  _IidXlnCO_3 |  -.5911753   .5030089    -1.18   0.241    -1.580485    .3981348
  _IidXlnCO_4 |  -.9062558   .5041282    -1.80   0.073    -1.897767    .0852558
  _IidXlnCO_5 |  -.2700905   .4803258    -0.56   0.574    -1.214788    .6746069
  _IidXlnCO_6 |  -.4817362   .4848666    -0.99   0.321    -1.435364    .4718918
  _IidXlnCO_7 |  -.3338866   .4944435    -0.68   0.500     -1.30635    .6385772
  _IidXlnCO_8 |  -1.034463   .4932714    -2.10   0.037    -2.004622   -.0643047
```

```
 _IidXlnCO_9 |  -.7298646    .538345    -1.36   0.176    -1.788673    .3289441
_IidXlnCO_10 |  -.6191171   .5185401    -1.19   0.233    -1.638974    .4007396
_IidXlnCO_11 |   -1.31699   .4849049    -2.72   0.007    -2.270693   -.3632863
_IidXlnCO_12 |  -.6970705   .5070844    -1.37   0.170    -1.694396    .3002552
_IidXlnCO_13 |  -.5330109   .4829867    -1.10   0.271    -1.482942    .4169199
_IidXlnCO_14 |  -.0275129   .4939434    -0.06   0.956    -.9989932    .9439674
_IidXlnCO_15 |   .8011863   .5050393     1.59   0.114    -.1921173     1.79449
_IidXlnCO_16 |    .566662   .4885764     1.16   0.247    -.3942626    1.527587
_IidXlnCO_17 |  -1.192849   .5074758    -2.35   0.019    -2.190945   -.1947538
_IidXlnCO_18 |  -1.915974   .4950568    -3.87   0.000    -2.889644    -.942304
_IidXlnCO_19 |  -2.522239   .4783071    -5.27   0.000    -3.462966   -1.581512
_IidXlnCO_20 |  -1.616987   .4885647    -3.31   0.001    -2.577888   -.6560852
_IidXlnCO_21 |  -.1668102   .4955661    -0.34   0.737    -1.141482    .8078616
  lnSKILLTL2 |  -.9837025   1.205216    -0.82   0.415    -3.354103    1.386698
 _IidXlnSK_2 |   .7067086   1.669968     0.42   0.672    -2.577759    3.991176
 _IidXlnSK_3 |   -1.44655   1.711311    -0.85   0.399     -4.81233     1.91923
 _IidXlnSK_4 |   1.451269   1.640473     0.88   0.377    -1.775188    4.677727
 _IidXlnSK_5 |   1.059546   1.618974     0.65   0.513    -2.124628    4.243719
 _IidXlnSK_6 |   1.081916   1.608482     0.67   0.502    -2.081621    4.245453
 _IidXlnSK_7 |   .2307377   1.608832     0.14   0.886    -2.933488    3.394964
 _IidXlnSK_8 |  -.3128141    1.67345    -0.19   0.852     -3.60413    2.978502
 _IidXlnSK_9 |   5.788814   1.766065     3.28   0.001     2.315345    9.262284
_IidXlnSK_10 |  -.4689307   1.644781    -0.29   0.776     -3.70386    2.765998
_IidXlnSK_11 |  -2.071723   1.622695    -1.28   0.203    -5.263214    1.119767
_IidXlnSK_12 |  -.8043671   1.701324    -0.47   0.637    -4.150504     2.54177
_IidXlnSK_13 |   2.915279   1.672478     1.74   0.082    -.3741239    6.204683
_IidXlnSK_14 |  -.8008423   1.617523    -0.50   0.621    -3.982161    2.380476
_IidXlnSK_15 |   3.723149   1.608521     2.31   0.021     .5595347    6.886763
_IidXlnSK_16 |  -1.490551   1.629515    -0.91   0.361    -4.695456    1.714355
_IidXlnSK_17 |  -3.045341   1.674933    -1.82   0.070    -6.339573    .2488903
_IidXlnSK_18 |  -.7024391   1.611543    -0.44   0.663    -3.871996    2.467118
_IidXlnSK_19 |  -.0560702   1.738304    -0.03   0.974     -3.47494    3.362799
_IidXlnSK_20 |   1.829652   1.651219     1.11   0.269     -1.41794    5.077245
_IidXlnSK_21 |   .8309436   1.623789     0.51   0.609    -2.362699    4.024586
   lnOPENTL2 |  -.8737359   1.056383    -0.83   0.409    -2.951414    1.203942
 _IidXlnOP_2 |  -1.829181   1.461732    -1.25   0.212    -4.704093    1.045731
 _IidXlnOP_3 |  -1.428909   1.491339    -0.96   0.339    -4.362051    1.504233
 _IidXlnOP_4 |   .3172468   1.422982     0.22   0.824    -2.481452    3.115946
 _IidXlnOP_5 |   3.098354   1.438376     2.15   0.032     .2693781    5.927329
 _IidXlnOP_6 |   .7788704    1.49481     0.52   0.603    -2.161099     3.71884
 _IidXlnOP_7 |  -.0357368   1.442575    -0.02   0.980     -2.87297    2.801497
 _IidXlnOP_8 |  -.0255781   1.445749    -0.02   0.986    -2.869054    2.817898
 _IidXlnOP_9 |   -2.29562    1.53103    -1.50   0.135    -5.306826    .7155863
_IidXlnOP_10 |   3.347031   1.427177     2.35   0.020     .5400805    6.153981
_IidXlnOP_11 |  -.7786004   1.475463    -0.53   0.598    -3.680518    2.123317
_IidXlnOP_12 |   1.055476   1.489593     0.71   0.479    -1.874233    3.985186
_IidXlnOP_13 |   2.273025   1.477361     1.54   0.125    -.6326259    5.178675
_IidXlnOP_14 |   3.392489   1.423743     2.38   0.018     .5922939    6.192685
_IidXlnOP_15 |  -.9081172   1.418757    -0.64   0.523    -3.698506    1.882271
_IidXlnOP_16 |   1.758822   1.432426     1.23   0.220    -1.058452    4.576097
_IidXlnOP_17 |    1.29799    1.45631     0.89   0.373    -1.566259    4.162238
_IidXlnOP_18 |  -1.754627   1.465417    -1.20   0.232    -4.636787    1.127533
_IidXlnOP_19 |   1.675465   1.482431     1.13   0.259    -1.240158    4.591088
_IidXlnOP_20 |  -1.046494   1.474658    -0.71   0.478    -3.946828     1.85384
_IidXlnOP_21 |   .9585003   1.433616     0.67   0.504    -1.861114    3.778115
     lnFERTL2 |   .0531045     .07624     0.70   0.487    -.0968431    .2030522
 _IidXlnFE_2 |   .1507181   .1097454     1.37   0.171    -.0651275    .3665637
 _IidXlnFE_3 |   .0369502   .1087565     0.34   0.734    -.1769504    .2508507
 _IidXlnFE_4 |  -.1152803   .1073704    -1.07   0.284    -.3264546    .0958941
 _IidXlnFE_5 |  -.2477329   .1085957    -2.28   0.023    -.4613172   -.0341486
 _IidXlnFE_6 |    .175484   .1096034     1.60   0.110    -.0400822    .3910502
 _IidXlnFE_7 |   .0738626    .108987     0.68   0.498    -.1404913    .2882164
 _IidXlnFE_8 |   .0111464   .1107759     0.10   0.920    -.2067259    .2290187
```

```
 _IidXlnFE_9 |  -.0693285   .1076073    -0.64   0.520    -.2809688    .1423118
_IidXlnFE_10 |  -.0351531   .1075604    -0.33   0.744    -.2467011     .176395
_IidXlnFE_11 |   .0211681   .1087819     0.19   0.846    -.1927826    .2351187
_IidXlnFE_12 |  -.0564374    .108994    -0.52   0.605    -.2708051    .1579303
_IidXlnFE_13 |  -.0081752   .1085299    -0.08   0.940    -.2216301    .2052798
_IidXlnFE_14 |   .0144987   .1080496     0.13   0.893    -.1980117    .2270091
_IidXlnFE_15 |   .1085933   .1079752     1.01   0.315    -.1037706    .3209573
_IidXlnFE_16 |  -.0057951    .108605    -0.05   0.957    -.2193977    .2078075
_IidXlnFE_17 |  -.0863754   .1094914    -0.79   0.431    -.3017214    .1289706
_IidXlnFE_18 |   .0173218   .1084975     0.16   0.873    -.1960695    .2307131
_IidXlnFE_19 |  -.4474465   .1110387    -4.03   0.000    -.6658357   -.2290573
_IidXlnFE_20 |  -.0354079    .108855    -0.33   0.745    -.2495022    .1786864
_IidXlnFE_21 |  -.0406549   .1074275    -0.38   0.705    -.2519416    .1706317
     lnTCTL2 |  -.6551358   1.118497    -0.59   0.558    -2.854979    1.544707
 _IidXlnTC_2 |   2.530407   1.573412     1.61   0.109    -.5641545    5.624969
 _IidXlnTC_3 |   .7323033   1.570311     0.47   0.641    -2.356159    3.820766
 _IidXlnTC_4 |   .0771453   1.554278     0.05   0.960    -2.979785    3.134076
 _IidXlnTC_5 |  -1.068715   1.532797    -0.70   0.486    -4.083397    1.945967
 _IidXlnTC_6 |   2.017755   1.555822     1.30   0.196    -1.042213    5.077722
 _IidXlnTC_7 |   .7059536   1.563723     0.45   0.652    -2.369553     3.78146
 _IidXlnTC_8 |   .8348708   1.579305     0.53   0.597    -2.271282    3.941024
 _IidXlnTC_9 |  -.1647896   1.621976    -0.10   0.919    -3.354867    3.025288
_IidXlnTC_10 |    .971268   1.569929     0.62   0.537    -2.116445    4.058981
_IidXlnTC_11 |   .3827959   1.546091     0.25   0.805    -2.658032    3.423624
_IidXlnTC_12 |   1.090268   1.554291     0.70   0.483    -1.966687    4.147223
_IidXlnTC_13 |   1.643676   1.570969     1.05   0.296    -1.446082    4.733434
_IidXlnTC_14 |   2.169522   1.539137     1.41   0.160    -.8576301    5.196673
_IidXlnTC_15 |  -1.598957   1.537865    -1.04   0.299    -4.623607    1.425692
_IidXlnTC_16 |  -3.090407   1.529921    -2.02   0.044    -6.099431   -.0813828
_IidXlnTC_17 |   1.901606   1.581536     1.20   0.230    -1.208934    5.012146
_IidXlnTC_18 |   2.237671   1.575155     1.42   0.156    -.8603202    5.335661
_IidXlnTC_19 |  -.2030253   1.564389    -0.13   0.897    -3.279841    2.873791
_IidXlnTC_20 |    .506868   1.568523     0.32   0.747    -2.578079    3.591815
_IidXlnTC_21 |   1.365406   1.556177     0.88   0.381    -1.695258     4.42607
   lnRISKTL2 |   1.150661    1.75828     0.65   0.513    -2.307498    4.608819
 _IidXlnRI_2 |   2.499549   2.526616     0.99   0.323    -2.469761    7.468859
 _IidXlnRI_3 |  -2.082396   2.493124    -0.84   0.404    -6.985834    2.821041
 _IidXlnRI_4 |   1.144597    2.47226     0.46   0.644    -3.717806       6.007
 _IidXlnRI_5 |   .1990562   2.427103     0.08   0.935    -4.574532    4.972645
 _IidXlnRI_6 |  -2.822244   2.492679    -1.13   0.258    -7.724807     2.08032
 _IidXlnRI_7 |  -2.172722    2.50419    -0.87   0.386    -7.097925    2.752481
 _IidXlnRI_8 |  -2.454418   2.532909    -0.97   0.333    -7.436105    2.527269
 _IidXlnRI_9 |   5.426003   2.475843     2.19   0.029     .5565537    10.29545
_IidXlnRI_10 |  -.1651862   2.588172    -0.06   0.949    -5.255562     4.92519
_IidXlnRI_11 |  -5.253176   2.480045    -2.12   0.035    -10.13089   -.3754614
_IidXlnRI_12 |  -5.408094     2.6618    -2.03   0.043    -10.64328   -.1729064
_IidXlnRI_13 |  -3.498098   2.492013    -1.40   0.161    -8.399351    1.403156
_IidXlnRI_14 |  -1.598148   2.488449    -0.64   0.521    -6.492391    3.296095
_IidXlnRI_15 |  -4.969254   2.460031    -2.02   0.044    -9.807605    -.130904
_IidXlnRI_16 |   .4780089   2.564217     0.19   0.852    -4.565254    5.521272
_IidXlnRI_17 |  -7.920547   2.580239    -3.07   0.002    -12.99532   -2.845773
_IidXlnRI_18 |  -.3090946   2.460567    -0.13   0.900    -5.148499     4.53031
_IidXlnRI_19 |  -2.540486   2.475178    -1.03   0.305    -7.408628    2.327656
_IidXlnRI_20 |   2.304269   2.466298     0.93   0.351    -2.546407    7.154945
_IidXlnRI_21 |  -.6061853   2.484125    -0.24   0.807    -5.491924    4.279554
       _cons |  -.0000725   4.156601    -0.00   1.000    -8.175212    8.175067
-------------+----------------------------------------------------------------
     sigma_u |  34.665942
     sigma_e |  .13319495
         rho |  .99998524   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(20, 349) =     3.08             Prob > F = 0.0000
```

```
//CCEMG
All coefficients represent averages across groups (group variable: id)
Coefficient averages computed as outlier-robust means (using rreg)
Mean Group type estimation                      Number of obs      =       567
Group variable: id                              Number of groups   =        21
                                                Obs per group: min =        27
                                                               avg =      27.0
                                                               max =        27
                                                Wald chi2(8)       =     16.80
                                                Prob > chi2        =    0.0322
------------------------------------------------------------------------------
       lnFDI |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     lnGDPL2 |   1.530559   .6716905     2.28   0.023     .2140697    2.847048
     lnTAXL2 |  -.3112115   .1932037    -1.61   0.107    -.6898839    .0674609
    lnCOSTL2 |   .2968496   .5412835     0.55   0.583    -.7640465    1.357746
   lnSKILLL2 |   .3834125   .4743651     0.81   0.419    -.5463261    1.313151
    lnOPENL2 |   .1236433   .5748871     0.22   0.830    -1.003115    1.250401
     lnFERL2 |  -.0300808   .0159609    -1.88   0.059    -.0613637    .0012021
      lnTCL2 |    .009604   .1031742     0.09   0.926    -.1926137    .2118217
    lnRISKL2 |   .7293028   .3455124     2.11   0.035     .0521109    1.406495
   lnFDI_avg |   1.009809   .2214485     4.56   0.000     .5757777     1.44384
 lnGDPL2_avg |  -1.748931   1.012901    -1.73   0.084    -3.734181    .2363197
 lnTAXL2_avg |   .4468965   .2200645     2.03   0.042      .015578     .878215
lnCOSTL2_avg |  -.3249281   .6087294    -0.53   0.593    -1.518016    .8681595
lnSKILLL2_~g |  -.3108531   .5716972    -0.54   0.587    -1.431359    .8096528
lnOPENL2_avg |   .2141695   .7291862     0.29   0.769    -1.215009    1.643348
 lnFERL2_avg |   .0577505   .0210774     2.74   0.006     .0164395    .0990614
  lnTCL2_avg |   .9233487   .2693962     3.43   0.001     .3953418    1.451356
lnRISKL2_avg |  -1.229952   .8865944    -1.39   0.165    -2.967645    .5077409
       _cons |   1.347003   9.392998     0.14   0.886    -17.06294    19.75694
------------------------------------------------------------------------------
Root Mean Squared Error (sigma): 0.0627
```

APPENDICES OF STUDY II

Appendix II.1:

List of countries in the sample 1

- Albania
- Argentina
- Armenia
- Australia
- Austria
- Bahrain
- Bangladesh
- Barbados
- Belgium
- Belize
- Benin
- Bolivia
- Botswana
- Brazil
- Brunei
- Bulgaria
- Burundi
- Cambodia
- Cameroon
- Canada
- Cen. African Rep.
- Chile
- China
- Colombia
- Congo, Dem. Rep.
- Congo, Republic of
- Costa Rica
- Cote d`Ivoire
- Croatia
- Cyprus
- Czech Republic
- Denmark
- Dominican Rep.
- Ecuador
- Egypt
- El Salvador
- Estonia
- Fiji
- Finland
- France
- Gabon
- Gambia, The
- Germany
- Ghana
- Greece
- Guatemala
- Honduras
- Hong Kong
- Hungary
- Iceland
- India
- Indonesia
- Iran
- Iraq
- Ireland
- Israel
- Italy
- Jamaica
- Japan
- Jordan
- Kazakhstan
- Kenya
- Korea
- Kuwait
- Kyrgyzstan
- Laos
- Latvia
- Lesotho
- Liberia
- Lithuania
- Luxembourg
- Macao
- Malawi
- Malaysia
- Maldives
- Mali
- Malta
- Mauritania
- Mauritius
- Mexico
- Moldova
- Mongolia
- Morocco
- Mozambique
- Namibia
- Nepal
- Netherlands
- New Zealand
- Niger
- Norway
- Pakistan
- Panama
- Paraguay
- Peru
- Philippines
- Poland
- Portugal
- Qatar
- Romania
- Russia
- Rwanda
- Saudi Arabia
- Senegal
- Serbia
- Sierra Leone
- Singapore
- Slovak Republic
- Slovenia
- South Africa
- Spain
- Sri Lanka
- Sudan
- Swaziland
- Sweden
- Switzerland
- Syria
- Taiwan
- Tajikistan
- Tanzania
- Thailand
- Togo
- Trinidad &Tobago
- Tunisia
- Turkey
- Uganda
- Ukraine
- United Kingdom
- United States
- Uruguay
- Venezuela
- Vietnam
- Yemen
- Zambia
- Zimbabwe

*Sample 2 consists of the countries in Sample 1 excepting Bahrain, Gabon, Iran, Iraq, Kuwait, Qatar and Saudi Arabia.

Appendix II.2:

List of sub-Saharan African countries

Burundi	Mauritania
Benin	Mauritius
Botswana	Malawi
Cen. African Rep.	Namibia
Cote d`Ivoire	Niger
Cameroon	Rwanda
Congo, Dem. Rep.	Sudan
Congo, Republic of	Senegal
Gabon	Sierra Leone
Ghana	Swaziland
Gambia, The	Togo
Kenya	Tanzania
Liberia	Uganda
Lesotho	South Africa
Mali	Zambia
Mozambique	Zimbabwe

Appendix II.3:

List of East Asian countries

Brunei	Macao
Cambodia	Malaysia
China	Mongolia
Fiji	Philippines
Hong Kong	Singapore
Indonesia	Taiwan
Korea	Thailand
Laos	Vietnam

Appendix II.4: Experiment instrument sets in the SGMM2R estimation – Sample 2

Instrument set	
1. 'Collapse' the standard instrument set (used in the SGMM2R regression in Table II.4.1)	Hansen (p-value): 0.01 Dif. Hansen (p-value): 0.01 No. of instr.(a): 47
2. Use $\ln(Y/L)_{i,t-2}$, $\ln s_{i,t-2}$, $\ln(n_{i,t-2}+g+\delta)$ and $\ln h_{i,t-1}$ and as instruments for first differenced equations; combined with $\Delta\ln(Y/L)_{i,t-1}$, $\Delta\ln s_{i,t-1}$, $\Delta\ln(n_{i,t-1}+g+\delta)$ and $\Delta\ln h_{it}$ as instruments for level equations.	Hansen (p-value): 0.03 Dif. Hansen (p-value): 0.57 No. of instr.(a): 85
3. Use $\ln(Y/L)_{i,t-2}$ and $\ln(Y/L)_{i,t-3}$, $\ln s_{i,t-2}$ and $\ln s_{i,t-3}$, $\ln(n_{i,t-2}+g+\delta)$ and $\ln(n_{i,t-3}+g+\delta)$, and $\ln h_{i,t-1}$ and $\ln h_{i,t-2}$ as instruments for first differenced equations; combined with $\Delta\ln(Y/L)_{i,t-1}$, $\Delta\ln s_{i,t-1}$, $\Delta\ln(n_{i,t-1}+g+\delta)$ and $\Delta\ln h_{it}$ as instruments for level equations.	Hansen (p-value): 0.33 Dif. Hansen (p-value): 0.99 No. of instr.(a): 123
4. Use $\ln(Y/L)_{i,t-2}$ ... $\ln(Y/L)_{i,t-4}$, $\ln s_{i,t-2}$ ... $\ln s_{i,t-4}$, $\ln(n_{i,t-2}+g+\delta)\ldots$ $\ln(n_{i,t-4}+g+\delta)$, and $\ln h_{i,t-1}$ … $\ln h_{i,t-3}$ as instruments for first differenced equations; combined with $\Delta\ln(Y/L)_{i,t-1}$, $\Delta\ln s_{i,t-1}$, $\Delta\ln(n_{i,t-1}+g+\delta)$ and $\Delta\ln h_{it}$ as instruments for level equations.	Hansen (p-value): 0.96 Dif. Hansen (p-value): 1.00 No. of instr.(a): 157
5. Use $\ln(Y/L)_{i,t-3}$, $\ln s_{i,t-3}$, $\ln(n_{i,t-3}+g+\delta)$ and $\ln h_{i,t-2}$ and as instruments for first differenced equations; combined with $\Delta\ln(Y/L)_{i,t-2}$, $\Delta\ln s_{i,t-2}$, $\Delta\ln(n_{i,t-2}+g+\delta)$ and $\Delta\ln h_{i,t-1}$ as instruments for level equations.	Hansen (p-value): 0.18 Dif. Hansen (p-value): 0.93 No. of instr.(a): 77
6. Use $\ln(Y/L)_{i,t-3}$ and $\ln(Y/L)_{i,t-4}$, $\ln s_{i,t-3}$ and $\ln s_{i,t-4}$, $\ln(n_{i,t-3}+g+\delta)$ and $\ln(n_{i,t-4}+g+\delta)$, and $\ln h_{i,t-2}$ and $\ln h_{i,t-3}$ as instruments for first differenced equations; combined with $\Delta\ln(Y/L)_{i,t-2}$, $\Delta\ln s_{i,t-2}$, $\Delta\ln(n_{i,t-2}+g+\delta)$ and $\Delta\ln h_{i,t-1}$ as instruments for level equations.	Hansen (p-value): 0.21 Dif. Hansen (p-value): 0.38 No. of instr.(a): 111
7. Use $\ln(Y/L)_{i,t-4}$, $\ln s_{i,t-4}$, $\ln(n_{i,t-4}+g+\delta)$ and $\ln h_{i,t-3}$ and as instruments for first differenced equations; combined with $\Delta\ln(Y/L)_{i,t-3}$, $\Delta\ln s_{i,t-3}$, $\Delta\ln(n_{i,t-3}+g+\delta)$ and $\Delta\ln h_{i,t-2}$ as instruments for level equations.	Hansen (p-value): 0.01 Dif. Hansen (p-value): 0.19 No. of instr.(a): 69
8. Use $\ln(Y/L)_{i,t-4}$ and $\ln(Y/L)_{i,t-5}$, $\ln s_{i,t-4}$ and $\ln s_{i,t-5}$, $\ln(n_{i,t-4}+g+\delta)$ and $\ln(n_{i,t-5}+g+\delta)$, and $\ln h_{i,t-3}$ … $\ln h_{i,t-4}$ as instruments for first differenced equations; combined with $\Delta\ln(Y/L)_{i,t-3}$, $\Delta\ln s_{i,t-3}$, $\Delta\ln(n_{i,t-3}+g+\delta)$ and $\Delta\ln h_{i,t-1}$ as instruments for level equations.	Hansen (p-value): 0.15 Dif. Hansen (p-value): 0.26 No. of instr.(a): 99

(a) The number of instruments used for endogenous and predetermined variables.

Appendix II.5: STATA output – POLS, WG, DGMM2R_SIS and SGMM2R_SIS regressions of the HCASM – Sample 1

```
//POLS
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
note: _Iperiod_3 omitted because of collinearity
Linear regression                                      Number of obs =    1170
                                                       F( 15,   133) = 2996.57
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9823
                                                       Root MSE      =  .16534

                                   (Std. Err. adjusted for 134 clusters in id)
------------------------------------------------------------------------------
             |               Robust
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9605751   .0081827   117.39   0.000       .94439    .9767603
             |
        ln_s |   .0729056   .0161574     4.51   0.000     .0409469    .1048643
      ln_ngd |  -.1046503   .0243927    -4.29   0.000     -.152898   -.0564025
        ln_h |   .0433543   .0166252     2.61   0.010     .0104703    .0762383
  _Iperiod_2 |   .0349899   .0136519     2.56   0.011      .007987    .0619929
  _Iperiod_3 |  (omitted)
  _Iperiod_4 |   .0430938   .0158413     2.72   0.007     .0117602    .0744273
  _Iperiod_5 |   .0491431   .0169058     2.91   0.004     .0157041     .082582
  _Iperiod_6 |  -.0159613   .0204681    -0.78   0.437    -.0564464    .0245237
  _Iperiod_7 |  -.0253981   .0194648    -1.30   0.194    -.0638986    .0131024
  _Iperiod_8 |  -.1355421   .0206745    -6.56   0.000    -.1764355   -.0946487
  _Iperiod_9 |  -.0777018   .0202064    -3.85   0.000    -.1176694   -.0377343
 _Iperiod_10 |  -.1516648   .0294149    -5.16   0.000    -.2098464   -.0934833
 _Iperiod_11 |  -.0507621   .0201399    -2.52   0.013     -.090598   -.0109262
 _Iperiod_12 |  -.0417071    .018531    -2.25   0.026    -.0783606   -.0050535
 _Iperiod_13 |  -.0638313   .0176018    -3.63   0.000     -.098647   -.0290156
       _cons |   .2804512   .1028983     2.73   0.007     .0769223    .4839802
------------------------------------------------------------------------------

//WG
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
note: _Iperiod_13 omitted because of collinearity
Fixed-effects (within) regression               Number of obs      =      1170
Group variable: id                              Number of groups   =       134
R-sq:  within  = 0.8507                         Obs per group: min =         3
       between = 0.9925                                        avg =       8.7
       overall = 0.9744                                        max =        12
                                                F(15,133)          =    379.71
corr(u_i, Xb)  = 0.8890                         Prob > F           =    0.0000

                                   (Std. Err. adjusted for 134 clusters in id)
------------------------------------------------------------------------------
             |               Robust
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .7697616   .0346472    22.22   0.000     .7012309    .8382924
             |
        ln_s |    .088309   .0247636     3.57   0.001     .0393276    .1372905
      ln_ngd |  -.1089469   .0304793    -3.57   0.000    -.1692338     -.04866
        ln_h |  -.0928876   .0355852    -2.61   0.010    -.1632736   -.0225015
  _Iperiod_2 |  -.2608081   .0660243    -3.95   0.000    -.3914016   -.1302146
  _Iperiod_3 |  -.2607078   .0571316    -4.56   0.000     -.373712   -.1477036
```

```
  _Iperiod_4 |  -.1894486    .052112    -3.64   0.000     -.292524   -.0863732
  _Iperiod_5 |  -.1356123   .0466465    -2.91   0.004    -.2278773   -.0433473
  _Iperiod_6 |  -.1420974   .0370872    -3.83   0.000    -.2154545   -.0687404
  _Iperiod_7 |  -.1137364   .0317811    -3.58   0.000    -.1765981   -.0508747
  _Iperiod_8 |  -.1843301   .0270859    -6.81   0.000     -.237905   -.1307551
  _Iperiod_9 |  -.1088204   .0245203    -4.44   0.000    -.1573206   -.0603202
 _Iperiod_10 |   -.147929   .0320148    -4.62   0.000    -.2112531   -.0846049
 _Iperiod_11 |  -.0413566   .0159542    -2.59   0.011    -.0729133   -.0097999
 _Iperiod_12 |   -.006229    .011463    -0.54   0.588    -.0289024    .0164445
 _Iperiod_13 |  (omitted)
       _cons |   2.373023   .4255527     5.58   0.000     1.531297     3.21475
-------------+----------------------------------------------------------------
     sigma_u |  .31316605
     sigma_e |  .14671756
         rho |  .82001489   (fraction of variance due to u_i)
------------------------------------------------------------------------------

//DGMM2R_SIS
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step difference GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1032
Time variable : period                          Number of groups   =       134
Number of instruments = 253                     Obs per group: min =         2
F(15, 134)    =    229.16                                      avg =      7.70
Prob > F      =     0.000                                      max =        11
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .7324607   .0587202    12.47   0.000     .6163225    .8485989
             |
        ln_s |   .0685941   .0288912     2.37   0.019     .0114523    .1257359
      ln_ngd |  -.0967773   .0360447    -2.68   0.008    -.1680674   -.0254872
        ln_h |  -.1226718   .0586989    -2.09   0.039    -.2387679   -.0065757
  _Iperiod_2 |  -.3379553   .1104347    -3.06   0.003    -.5563759   -.1195347
  _Iperiod_3 |  -.3267823   .0969929    -3.37   0.001    -.5186175   -.1349472
  _Iperiod_4 |  -.2538011   .0882821    -2.87   0.005    -.4284078   -.0791945
  _Iperiod_5 |  -.1870571   .0746251    -2.51   0.013    -.3346526   -.0394617
  _Iperiod_6 |   -.182111   .0587603    -3.10   0.002    -.2983287   -.0658934
  _Iperiod_7 |  -.1448327   .0468474    -3.09   0.002    -.2374886   -.0521768
  _Iperiod_8 |  -.2085009   .0380953    -5.47   0.000    -.2838468    -.133155
  _Iperiod_9 |   -.131116   .0337632    -3.88   0.000    -.1978937   -.0643383
 _Iperiod_10 |  -.1591534   .0366713    -4.34   0.000    -.2316828    -.086624
 _Iperiod_11 |  -.0552627   .0222586    -2.48   0.014    -.0992864   -.0112389
 _Iperiod_12 |  -.0146281   .0133265    -1.10   0.274    -.0409855    .0117294
------------------------------------------------------------------------------
Instruments for first differences equation
  Standard
    D.(_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13)
```

```
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(1/12).ln_h
    L(2/12).(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.37  Pr > z =  0.018
Arellano-Bond test for AR(2) in first differences: z =   0.89  Pr > z =  0.371
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(238)  = 508.13  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(238)  = 128.44  Prob > chi2 =  1.000
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  gmm(ln_y ln_s ln_ngd, lag(2 .))
    Hansen test excluding group:     chi2(62)   =  74.96  Prob > chi2 =  0.125
    Difference (null H = exogenous): chi2(176)  =  53.48  Prob > chi2 =  1.000
  gmm(ln_h, lag(1 .))
    Hansen test excluding group:     chi2(172)  = 127.57  Prob > chi2 =  0.995
    Difference (null H = exogenous): chi2(66)   =   0.87  Prob > chi2 =  1.000
  iv(_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13)
    Hansen test excluding group:     chi2(227)  = 126.02  Prob > chi2 =  1.000
    Difference (null H = exogenous): chi2(11)   =   2.42  Prob > chi2 =  0.996

//SGMM2R_SIS
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space , perm.
_Iperiod_13 dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 296                     Obs per group: min =         3
F(15, 133)    =   1097.08                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9433275     .01327    71.09   0.000     .9170799     .969575
             |
        ln_s |   .1056202   .0245866     4.30   0.000     .0569889    .1542515
      ln_ngd |  -.1491044    .040644    -3.67   0.000    -.2294966   -.0687123
        ln_h |   .0578138   .0240344     2.41   0.018     .0102747    .1053528
  _Iperiod_2 |   .0946095   .0292555     3.23   0.002     .0367432    .1524758
  _Iperiod_3 |   .0692253   .0231122     3.00   0.003     .0235103    .1149402
  _Iperiod_4 |   .1146171   .0232781     4.92   0.000      .068574    .1606603
  _Iperiod_5 |   .1218689    .022931     5.31   0.000     .0765122    .1672257
  _Iperiod_6 |   .0604986   .0214596     2.82   0.006     .0180524    .1029448
  _Iperiod_7 |   .0476186   .0201881     2.36   0.020     .0076872    .0875499
  _Iperiod_8 |  -.0617109    .020129    -3.07   0.003    -.1015253   -.0218966
  _Iperiod_9 |   .0028044   .0172391     0.16   0.871    -.0312938    .0369026
 _Iperiod_10 |  -.0763384   .0268872    -2.84   0.005    -.1295202   -.0231565
 _Iperiod_11 |   .0164191   .0154579     1.06   0.290     -.014156    .0469942
```

```
 _Iperiod_12 |   .0244051   .0118223     2.06   0.041      .001021    .0477892
       _cons |    .285229   .1434748     1.99   0.049     .0014414    .5690165
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(1/12).ln_h
    L(2/12).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    D.ln_h
    DL.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.53  Pr > z =  0.011
Arellano-Bond test for AR(2) in first differences: z =   0.78  Pr > z =  0.437
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(280)  = 605.53  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(280)  = 125.92  Prob > chi2 =  1.000
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(238)  = 126.91  Prob > chi2 =  1.000
    Difference (null H = exogenous): chi2(42)   =  -0.98  Prob > chi2 =  1.000
  gmm(ln_y ln_s ln_ngd, lag(2 .))
    Hansen test excluding group:     chi2(73)   =  88.14  Prob > chi2 =  0.109
    Difference (null H = exogenous): chi2(207)  =  37.79  Prob > chi2 =  1.000
  gmm(ln_h, lag(1 .))
    Hansen test excluding group:     chi2(203)  = 126.41  Prob > chi2 =  1.000
    Difference (null H = exogenous): chi2(77)   =  -0.49  Prob > chi2 =  1.000
  iv(_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(269)  = 128.08  Prob > chi2 =  1.000
    Difference (null H = exogenous): chi2(11)   =  -2.15  Prob > chi2 =  1.000
```

Appendix II.6: STATA output – DGMM2R_IS4, SGMM2R_IS3, SGMM2R_IS6, LSDVC1 and LSDVC2 regressions of the HCASM with valid instrument sets – Sample1

```
//DGMM2R_IS4
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step difference GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1032
Time variable : period                          Number of groups   =       134
Number of instruments = 125                     Obs per group: min =         2
F(15, 134)    =    122.49                                      avg =      7.70
Prob > F      =     0.000                                      max =        11
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .6133362   .0782459     7.84   0.000     .4585794    .7680931
             |
        ln_s |    .051146   .0304058     1.68   0.095    -.0089913    .1112834
      ln_ngd |   -.040396   .0632761    -0.64   0.524    -.1655452    .0847532
        ln_h |  -.2118653    .080197    -2.64   0.009     -.370481   -.0532497
  _Iperiod_2 |  -.5716163   .1609657    -3.55   0.001    -.8899785   -.2532541
  _Iperiod_3 |  -.5251091   .1437565    -3.65   0.000    -.8094344   -.2407838
  _Iperiod_4 |  -.4322337   .1299299    -3.33   0.001    -.6892124    -.175255
  _Iperiod_5 |  -.3342558   .1102315    -3.03   0.003    -.5522745   -.1162372
  _Iperiod_6 |  -.2973868   .0879011    -3.38   0.001      -.47124   -.1235337
  _Iperiod_7 |  -.2394053   .0708336    -3.38   0.001    -.3795018   -.0993087
  _Iperiod_8 |  -.2769032   .0535574    -5.17   0.000    -.3828303    -.170976
  _Iperiod_9 |  -.1993948   .0481189    -4.14   0.000    -.2945657   -.1042239
 _Iperiod_10 |  -.1953709   .0456731    -4.28   0.000    -.2857044   -.1050374
 _Iperiod_11 |  -.0908421    .028704    -3.16   0.002    -.1476135   -.0340707
 _Iperiod_12 |  -.0300455   .0154111    -1.95   0.053    -.0605259     .000435
------------------------------------------------------------------------------
Instruments for first differences equation
  Standard
    D.(_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13)
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(1/3).ln_h
    L(2/4).(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.20  Pr > z =  0.028
Arellano-Bond test for AR(2) in first differences: z =   1.02  Pr > z =  0.307
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(110)  = 388.04  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(110)  = 116.77  Prob > chi2 =  0.311
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
```

```
  gmm(ln_y ln_s ln_ngd, lag(2 4))
    Hansen test excluding group:     chi2(26)    =  37.66  Prob > chi2 =  0.065
    Difference (null H = exogenous): chi2(84)    =  79.11  Prob > chi2 =  0.631
  gmm(ln_h, lag(1 3))
    Hansen test excluding group:     chi2(80)    = 102.15  Prob > chi2 =  0.048
    Difference (null H = exogenous): chi2(30)    =  14.62  Prob > chi2 =  0.992
  iv(_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13)
    Hansen test excluding group:     chi2(99)    = 100.53  Prob > chi2 =  0.438
    Difference (null H = exogenous): chi2(11)    =  16.24  Prob > chi2 =  0.132

//SGMM2R_IS3
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 134                     Obs per group: min =         3
F(15, 133)    =   1039.51                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9491583   .0169551    55.98   0.000     .9156218    .9826949
             |
        ln_s |   .1108652   .0302577     3.66   0.000     .0510166    .1707139
      ln_ngd |  -.2072185   .0623745    -3.32   0.001    -.3305929   -.0838441
        ln_h |    .033007   .0285858     1.15   0.250    -.0235345    .0895486
  _Iperiod_2 |   .0711818   .0274109     2.60   0.010     .0169641    .1253996
  _Iperiod_3 |   .0386761   .0225293     1.72   0.088     -.005886    .0832382
  _Iperiod_4 |   .0961565   .0208097     4.62   0.000     .0549957    .1373172
  _Iperiod_5 |   .1039037    .023427     4.44   0.000     .0575661    .1502413
  _Iperiod_6 |   .0458461   .0226728     2.02   0.045     .0010003     .090692
  _Iperiod_7 |   .0403738   .0222737     1.81   0.072    -.0036828    .0844303
  _Iperiod_8 |  -.0751275   .0180514    -4.16   0.000    -.1108326   -.0394225
  _Iperiod_9 |  -.0025593   .0173546    -0.15   0.883     -.036886    .0317674
 _Iperiod_10 |  -.0887334    .026467    -3.35   0.001    -.1410841   -.0363826
 _Iperiod_11 |   .0140334   .0167751     0.84   0.404    -.0191471    .0472139
 _Iperiod_12 |   .0181871   .0149029     1.22   0.224    -.0112902    .0476644
       _cons |   .1318925   .1570283     0.84   0.402    -.1787034    .4424883
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(1/2).ln_h
    L(2/3).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
```

```
    D.ln_h
    DL.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.55  Pr > z =  0.011
Arellano-Bond test for AR(2) in first differences: z =   0.77  Pr > z =  0.442
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(118)  = 435.57  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(118)  = 126.91  Prob > chi2 =  0.271
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(76)   =  94.00  Prob > chi2 =  0.079
    Difference (null H = exogenous): chi2(42)   =  32.91  Prob > chi2 =  0.841
  gmm(ln_y ln_s ln_ngd, lag(2 3))
    Hansen test excluding group:     chi2(28)   =  40.36  Prob > chi2 =  0.061
    Difference (null H = exogenous): chi2(90)   =  86.55  Prob > chi2 =  0.583
  gmm(ln_h, lag(1 2))
    Hansen test excluding group:     chi2(86)   = 104.85  Prob > chi2 =  0.082
    Difference (null H = exogenous): chi2(32)   =  22.05  Prob > chi2 =  0.906
  iv(_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(107)  = 115.10  Prob > chi2 =  0.279
    Difference (null H = exogenous): chi2(11)   =  11.81  Prob > chi2 =  0.378

//SGMM2R_IS6
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 122                     Obs per group: min =         3
F(15, 133)    =   1290.55                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9693327   .0153923    62.98   0.000     .9388873    .9997781
             |
        ln_s |   .0878797   .0283445     3.10   0.002     .0318153     .143944
      ln_ngd |  -.1341353   .0707267    -1.90   0.060    -.2740299    .0057594
        ln_h |   .0295511   .0348958     0.85   0.399    -.0394714    .0985737
  _Iperiod_2 |   .0968734   .0318583     3.04   0.003     .0338589    .1598879
  _Iperiod_3 |   .0571315   .0295933     1.93   0.056    -.0014029    .1156658
  _Iperiod_4 |   .1116336   .0264772     4.22   0.000     .0592626    .1640045
  _Iperiod_5 |   .1188615   .0279948     4.25   0.000     .0634888    .1742342
  _Iperiod_6 |   .0499449   .0212506     2.35   0.020      .007912    .0919778
  _Iperiod_7 |   .0366961   .0208173     1.76   0.080    -.0044796    .0778719
  _Iperiod_8 |  -.0676758    .020985    -3.22   0.002    -.1091832   -.0261683
  _Iperiod_9 |   .0007472   .0192598     0.04   0.969    -.0373479    .0388423
```

```
 _Iperiod_10 |  -.0662208    .025535    -2.59   0.011     -.116728   -.0157136
 _Iperiod_11 |   .0199016   .0184698     1.08   0.283     -.016631    .0564342
 _Iperiod_12 |   .0238217   .0134517     1.77   0.079    -.0027852    .0504287
       _cons |   .0974066   .1690253     0.58   0.565    -.2369189    .4317322
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.51  Pr > z =  0.012
Arellano-Bond test for AR(2) in first differences: z =   0.78  Pr > z =  0.438
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 350.80  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  = 114.27  Prob > chi2 =  0.274
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)   =  75.55  Prob > chi2 =  0.222
    Difference (null H = exogenous): chi2(39)   =  38.72  Prob > chi2 =  0.482
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)   =  41.26  Prob > chi2 =  0.029
    Difference (null H = exogenous): chi2(80)   =  73.01  Prob > chi2 =  0.697
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)   =  95.52  Prob > chi2 =  0.065
    Difference (null H = exogenous): chi2(30)   =  18.76  Prob > chi2 =  0.945
  iv(_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(95)   = 103.58  Prob > chi2 =  0.257
    Difference (null H = exogenous): chi2(11)   =  10.70  Prob > chi2 =  0.469

//LSDVC1
Note: Bias correction initialized by matrix my
note: Bias correction up to order O(1/T)
LSDVC dynamic regression
(bootstrapped SE)
------------------------------------------------------------------------------
        ln_y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |    .997312     .02138    46.65   0.000     .9554081    1.039216
             |
        ln_s |   .0822984   .0342961     2.40   0.016     .0150794    .1495175
      ln_ngd |  -.1194262   .0377542    -3.16   0.002    -.1934231   -.0454293
        ln_h |  -.0327408   .0694815    -0.47   0.637    -.1689221    .1034405
          p2 |   .0619562   .1003738     0.62   0.537    -.1347727    .2586852
          p3 |   .0237736    .093518     0.25   0.799    -.1595183    .2070655
          p4 |   .0537474   .0835714     0.64   0.520    -.1100494    .2175443
          p5 |   .0618511   .0763734     0.81   0.418    -.0878381    .2115403
          p6 |   .0009461   .0683759     0.01   0.989    -.1330682    .1349604
          p7 |  -.0024849   .0613415    -0.04   0.968    -.1227121    .1177423
          p8 |  -.1076069   .0533309    -2.02   0.044    -.2121335   -.0030802
          p9 |  -.0384971    .048259    -0.80   0.425     -.133083    .0560888
```

```
         p10 |  -.0955359    .045616    -2.09   0.036    -.1849416   -.0061303
         p11 |   .0097319   .0432239     0.23   0.822    -.0749854    .0944492
         p12 |    .020205   .0404278     0.50   0.617    -.0590319     .099442
------------------------------------------------------------------------------

//LSDVC2
Note: Bias correction initialized by matrix my
note: Bias correction up to order O(1/T)
LSDVC dynamic regression
(bootstrapped SE)
------------------------------------------------------------------------------
        ln_y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9868624   .0205347    48.06   0.000     .9466152     1.02711
             |
        ln_s |    .082141   .0293006     2.80   0.005     .0247128    .1395691
      ln_ngd |  -.1129759   .0325365    -3.47   0.001    -.1767462   -.0492055
        ln_h |  -.0385142   .0590368    -0.65   0.514    -.1542242    .0771958
          p2 |    .043497   .0865839     0.50   0.615    -.1262043    .2131982
          p3 |   .0074993   .0805058     0.09   0.926    -.1502891    .1652877
          p4 |   .0410252   .0719167     0.57   0.568    -.0999289    .1819793
          p5 |   .0510734   .0657421     0.78   0.437    -.0777789    .1799256
          p6 |  -.0073622   .0587066    -0.13   0.900     -.122425    .1077006
          p7 |   -.008722   .0526704    -0.17   0.868     -.111954      .09451
          p8 |  -.1098026   .0458666    -2.39   0.017    -.1996995   -.0199057
          p9 |  -.0406154   .0415691    -0.98   0.329    -.1220893    .0408585
         p10 |  -.0966343   .0393747    -2.45   0.014    -.1738073   -.0194612
         p11 |   .0059287   .0374147     0.16   0.874    -.0674027    .0792602
         p12 |   .0182992    .034995     0.52   0.601    -.0502899    .0868882
------------------------------------------------------------------------------
```

Appendix II.7: STATA output – SGMM2R_IS6 regression of the HCASM – Sample 2

```
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1114
Time variable : period                          Number of groups   =       127
Number of instruments = 122                     Obs per group: min =         3
F(15, 126)    =   1332.85                                      avg =      8.77
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9875377   .0143357    68.89   0.000     .9591677    1.015908
             |
        ln_s |    .075429   .0286363     2.63   0.009     .0187586    .1320993
      ln_ngd |   .0263655   .0501575     0.53   0.600    -.0728947    .1256257
        ln_h |    .042071    .036585     1.15   0.252    -.0303297    .1144716
  _Iperiod_2 |   .1178152   .0321766     3.66   0.000     .0541387    .1814916
  _Iperiod_3 |   .0763122   .0272835     2.80   0.006      .022319    .1303054
  _Iperiod_4 |   .1160123   .0263936     4.40   0.000     .0637802    .1682443
  _Iperiod_5 |   .1247738   .0313029     3.99   0.000     .0628263    .1867212
  _Iperiod_6 |   .0411138   .0224752     1.83   0.070     -.003364    .0855915
  _Iperiod_7 |   .0296694   .0234618     1.26   0.208    -.0167609    .0760996
  _Iperiod_8 |  -.0545706   .0206072    -2.65   0.009    -.0953517   -.0137896
  _Iperiod_9 |  -.0173998    .019459    -0.89   0.373    -.0559086    .0211089
 _Iperiod_10 |  -.0736548   .0257082    -2.87   0.005    -.1245305   -.0227791
 _Iperiod_11 |   .0059251   .0166139     0.36   0.722    -.0269534    .0388036
 _Iperiod_12 |   .0254468   .0139933     1.82   0.071    -.0022456    .0531392
       _cons |   .3258396   .1587474     2.05   0.042     .0116832    .6399961
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.34  Pr > z =  0.020
Arellano-Bond test for AR(2) in first differences: z =  -0.79  Pr > z =  0.430
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 324.28  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
```

```
Hansen test of overid. restrictions: chi2(106)  = 117.75  Prob > chi2 =  0.205
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)   =  76.66  Prob > chi2 =  0.196
    Difference (null H = exogenous): chi2(39)   =  41.09  Prob > chi2 =  0.379
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)   =  41.47  Prob > chi2 =  0.028
    Difference (null H = exogenous): chi2(80)   =  76.27  Prob > chi2 =  0.597
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)   =  99.71  Prob > chi2 =  0.035
    Difference (null H = exogenous): chi2(30)   =  18.04  Prob > chi2 =  0.958
  iv(_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(95)   = 109.94  Prob > chi2 =  0.140
    Difference (null H = exogenous): chi2(11)   =   7.81  Prob > chi2 =  0.730
```

Appendix II.8: STATA output – SGMM2R_IS3 and SGMM2R_IS6 estimation of the HCASM with SSA and EA dummies – Sample 1

```
//SGMM2R_IS3
/*The HCASM with SSA dummy*/
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 135                     Obs per group: min =         3
F(16, 133)    =   1134.86                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |    .929482   .0177386    52.40   0.000     .8943958    .9645681
             |
        ln_s |   .0965439   .0294511     3.28   0.001     .0382908     .154797
      ln_ngd |  -.1711863   .0651686    -2.63   0.010    -.3000872   -.0422853
        ln_h |   .0360768   .0260487     1.38   0.168    -.0154465    .0876001
         ssa |  -.0917296   .0280391    -3.27   0.001    -.1471898   -.0362695
  _Iperiod_2 |     .05554   .0252788     2.20   0.030     .0055396    .1055404
  _Iperiod_3 |   .0247642   .0228845     1.08   0.281    -.0205005    .0700289
  _Iperiod_4 |   .0802085    .022836     3.51   0.001     .0350397    .1253772
  _Iperiod_5 |   .0912473   .0232816     3.92   0.000     .0451971    .1372975
  _Iperiod_6 |   .0335059   .0209286     1.60   0.112    -.0078901    .0749018
  _Iperiod_7 |   .0297078   .0217006     1.37   0.173    -.0132152    .0726308
  _Iperiod_8 |  -.0775764   .0188896    -4.11   0.000    -.1149392   -.0402136
```

```
  _Iperiod_9 |  -.0092228   .0175349    -0.53   0.600    -.0439061    .0254606
 _Iperiod_10 |  -.0917921   .0265909    -3.45   0.001    -.1443879   -.0391964
 _Iperiod_11 |   .0093868   .0155426     0.60   0.547    -.0213559    .0401294
 _Iperiod_12 |   .0169676   .0145968     1.16   0.247    -.0119044    .0458395
       _cons |   .4122253   .1999938     2.06   0.041     .0166453    .8078052
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(1/2).ln_h
    L(2/3).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ssa _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    D.ln_h
    DL.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.54  Pr > z =  0.011
Arellano-Bond test for AR(2) in first differences: z =   0.77  Pr > z =  0.441
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(118)  = 441.05  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(118)  = 128.94  Prob > chi2 =  0.231
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(76)   =  94.63  Prob > chi2 =  0.073
    Difference (null H = exogenous): chi2(42)   =  34.32  Prob > chi2 =  0.794
  gmm(ln_y ln_s ln_ngd, lag(2 3))
    Hansen test excluding group:     chi2(28)   =  42.00  Prob > chi2 =  0.043
    Difference (null H = exogenous): chi2(90)   =  86.95  Prob > chi2 =  0.571
  gmm(ln_h, lag(1 2))
    Hansen test excluding group:     chi2(86)   = 107.43  Prob > chi2 =  0.059
    Difference (null H = exogenous): chi2(32)   =  21.51  Prob > chi2 =  0.920
  iv(ssa _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(106)  = 117.23  Prob > chi2 =  0.215
    Difference (null H = exogenous): chi2(12)   =  11.72  Prob > chi2 =  0.469

/*The HCASM with EA dummy*/
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 135                     Obs per group: min =         3
F(16, 133)    =   1053.77                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
```

```
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9567263   .0168782    56.68   0.000     .9233419    .9901107
             |
        ln_s |   .1001575   .0299246     3.35   0.001     .0409678    .1593472
      ln_ngd |  -.2195427   .0630245    -3.48   0.001    -.3442026   -.0948827
        ln_h |   .0229296   .0289174     0.79   0.429     -.034268    .0801271
          ea |   .0647104   .0260391     2.49   0.014     .0132061    .1162146
  _Iperiod_2 |   .0767372   .0268392     2.86   0.005     .0236503     .129824
  _Iperiod_3 |   .0392211   .0224265     1.75   0.083    -.0051376    .0835798
  _Iperiod_4 |   .0908212   .0207663     4.37   0.000     .0497464    .1318961
  _Iperiod_5 |   .0990942   .0240864     4.11   0.000     .0514522    .1467362
  _Iperiod_6 |   .0438772   .0224017     1.96   0.052    -.0004325    .0881868
  _Iperiod_7 |    .037681   .0217402     1.73   0.085    -.0053202    .0806822
  _Iperiod_8 |   -.077437   .0184685    -4.19   0.000     -.113967    -.040907
  _Iperiod_9 |  -.0047241    .017531    -0.27   0.788    -.0393999    .0299516
 _Iperiod_10 |  -.0899499   .0271746    -3.31   0.001    -.1437002   -.0361996
 _Iperiod_11 |   .0124936   .0181104     0.69   0.491     -.023328    .0483153
 _Iperiod_12 |   .0173034   .0148761     1.16   0.247     -.012121    .0467278
       _cons |    .020034   .1666358     0.12   0.904    -.3095653    .3496332
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(1/2).ln_h
    L(2/3).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ea _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    D.ln_h
    DL.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.55  Pr > z =  0.011
Arellano-Bond test for AR(2) in first differences: z =   0.75  Pr > z =  0.454
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(118)  = 420.09  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(118)  = 126.77  Prob > chi2 =  0.274
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(76)   =  93.60  Prob > chi2 =  0.083
    Difference (null H = exogenous): chi2(42)   =  33.17  Prob > chi2 =  0.833
  gmm(ln_y ln_s ln_ngd, lag(2 3))
    Hansen test excluding group:     chi2(28)   =  41.11  Prob > chi2 =  0.053
    Difference (null H = exogenous): chi2(90)   =  85.66  Prob > chi2 =  0.610
  gmm(ln_h, lag(1 2))
    Hansen test excluding group:     chi2(86)   = 105.29  Prob > chi2 =  0.077
    Difference (null H = exogenous): chi2(32)   =  21.48  Prob > chi2 =  0.921
  iv(ea _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_1 1 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(106)  = 114.78  Prob > chi2 =  0.263
    Difference (null H = exogenous): chi2(12)   =  11.99  Prob > chi2 =  0.447

/*The HCASM with SSA and EA dummies*/
```

Note: The number of instruments is reported in the STATA output is the total number of instruments used for all endogenous, predetermined and exogenous variables.

```
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 136                     Obs per group: min =         3
F(17, 133)    =   1276.39                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9411837   .0174791    53.85   0.000     .9066107    .9757568
             |
        ln_s |   .0879806   .0286246     3.07   0.003     .0313622     .144599
      ln_ngd |   -.191233   .0660571    -2.89   0.004    -.3218915   -.0605745
        ln_h |   .0233172   .0253157     0.92   0.359    -.0267564    .0733907
         ssa |  -.0757778   .0286534    -2.64   0.009    -.1324531   -.0191025
          ea |   .0442505   .0286759     1.54   0.125    -.0124692    .1009702
  _Iperiod_2 |   .0571566   .0265594     2.15   0.033     .0046231    .1096901
  _Iperiod_3 |    .023791   .0236016     1.01   0.315    -.0228919     .070474
  _Iperiod_4 |   .0740448   .0233831     3.17   0.002     .0277939    .1202957
  _Iperiod_5 |   .0860117   .0236328     3.64   0.000      .039267    .1327564
  _Iperiod_6 |   .0311765   .0212445     1.47   0.145    -.0108444    .0731973
  _Iperiod_7 |   .0274549   .0204446     1.34   0.182    -.0129837    .0678936
  _Iperiod_8 |  -.0809207   .0189014    -4.28   0.000    -.1183069   -.0435346
  _Iperiod_9 |  -.0108779   .0172589    -0.63   0.530    -.0450152    .0232595
 _Iperiod_10 |   -.092518   .0270019    -3.43   0.001    -.1459267   -.0391093
 _Iperiod_11 |   .0070704   .0171615     0.41   0.681    -.0268744    .0410153
 _Iperiod_12 |   .0154726   .0145023     1.07   0.288    -.0132124    .0441576
       _cons |   .2474279   .2057178     1.20   0.231    -.1594739    .6543297
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(1/2).ln_h
    L(2/3).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ssa ea _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    D.ln_h
    DL.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.53  Pr > z =  0.011
Arellano-Bond test for AR(2) in first differences: z =   0.75  Pr > z =  0.453
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(118)  = 428.56  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(118)  = 126.35  Prob > chi2 =  0.283
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
```

```
    Hansen test excluding group:     chi2(76)   =  94.61  Prob > chi2 =  0.073
    Difference (null H = exogenous): chi2(42)   =  31.74  Prob > chi2 =  0.875
  gmm(ln_y ln_s ln_ngd, lag(2 3))
    Hansen test excluding group:     chi2(28)   =  42.46  Prob > chi2 =  0.039
    Difference (null H = exogenous): chi2(90)   =  83.89  Prob > chi2 =  0.661
  gmm(ln_h, lag(1 2))
    Hansen test excluding group:     chi2(86)   = 105.94  Prob > chi2 =  0.071
    Difference (null H = exogenous): chi2(32)   =  20.42  Prob > chi2 =  0.944
  iv(ssa ea _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(105)  = 114.05  Prob > chi2 =  0.257
    Difference (null H = exogenous): chi2(13)   =  12.31  Prob > chi2 =  0.503

//SGMM2R_IS6
/*The HCASM with SSA dummy*/
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 123                     Obs per group: min =         3
F(16, 133)    =   1328.29                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9489542   .0165971    57.18   0.000     .9161257    .9817827
             |
        ln_s |   .0811618   .0288248     2.82   0.006     .0241476    .1381761
      ln_ngd |  -.1087015    .077594    -1.40   0.164    -.2621795    .0447765
        ln_h |   .0376633   .0336118     1.12   0.265    -.0288194    .1041461
         ssa |  -.0698638    .024028    -2.91   0.004    -.1173903   -.0223373
  _Iperiod_2 |   .0796345   .0307412     2.59   0.011     .0188295    .1404394
  _Iperiod_3 |   .0425341   .0282895     1.50   0.135    -.0134215    .0984896
  _Iperiod_4 |   .0976322   .0259115     3.77   0.000     .0463803     .148884
  _Iperiod_5 |   .1058238   .0279458     3.79   0.000     .0505481    .1610996
  _Iperiod_6 |    .040208    .021453     1.87   0.063    -.0022253    .0826412
  _Iperiod_7 |   .0276602   .0207206     1.33   0.184    -.0133244    .0686448
  _Iperiod_8 |  -.0705111   .0204377    -3.45   0.001     -.110936   -.0300861
  _Iperiod_9 |  -.0072858   .0205872    -0.35   0.724    -.0480064    .0334349
 _Iperiod_10 |  -.0707815   .0275861    -2.57   0.011    -.1253457   -.0162173
 _Iperiod_11 |   .0162834   .0189334     0.86   0.391    -.0211662     .053733
 _Iperiod_12 |   .0236177   .0130865     1.80   0.073    -.0022669    .0495022
       _cons |   .3549214   .2233507     1.59   0.114    -.0868576    .7967004
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
```

```
  Standard
    ssa _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.52  Pr > z =  0.012
Arellano-Bond test for AR(2) in first differences: z =   0.77  Pr > z =  0.441
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 351.04  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  = 115.72  Prob > chi2 =  0.244
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)   =  75.90  Prob > chi2 =  0.214
    Difference (null H = exogenous): chi2(39)   =  39.82  Prob > chi2 =  0.433
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)   =  41.56  Prob > chi2 =  0.027
    Difference (null H = exogenous): chi2(80)   =  74.16  Prob > chi2 =  0.663
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)   =  92.53  Prob > chi2 =  0.096
    Difference (null H = exogenous): chi2(30)   =  23.19  Prob > chi2 =  0.807
  iv(ssa _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(94)   = 105.26  Prob > chi2 =  0.201
    Difference (null H = exogenous): chi2(12)   =  10.46  Prob > chi2 =  0.576

/*The HCASM with EA dummy*/
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 123                     Obs per group: min =         3
F(16, 133)    =   1272.42                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9692846   .0145949    66.41   0.000     .9404164    .9981528
             |
        ln_s |   .0793158   .0266156     2.98   0.003     .0266712    .1319604
      ln_ngd |  -.1430341   .0693159    -2.06   0.041    -.2801382   -.0059301
        ln_h |   .0297126    .032082     0.93   0.356    -.0337444    .0931697
          ea |   .0677298   .0217709     3.11   0.002     .0246678    .1107918
  _Iperiod_2 |    .098714    .032187     3.07   0.003     .0350495    .1623786
  _Iperiod_3 |   .0595198   .0285325     2.09   0.039     .0030836    .1159559
```

```
  _Iperiod_4 |   .1095891   .0253987     4.31   0.000     .0593516    .1598267
  _Iperiod_5 |   .1155259    .028694     4.03   0.000     .0587703    .1722815
  _Iperiod_6 |   .0479797   .0217407     2.21   0.029     .0049774    .0909819
  _Iperiod_7 |   .0360528   .0208777     1.73   0.087    -.0052426    .0773481
  _Iperiod_8 |  -.0696327   .0211108    -3.30   0.001    -.1113891   -.0278764
  _Iperiod_9 |  -.0009688    .019786    -0.05   0.961    -.0401048    .0381671
 _Iperiod_10 |  -.0702431   .0256444    -2.74   0.007    -.1209668   -.0195194
 _Iperiod_11 |    .017083   .0179606     0.95   0.343    -.0184423    .0526083
 _Iperiod_12 |   .0218215   .0131614     1.66   0.100    -.0042113    .0478544
       _cons |   .0531623   .1736876     0.31   0.760    -.2903851    .3967096
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ea _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.50  Pr > z =  0.012
Arellano-Bond test for AR(2) in first differences: z =   0.77  Pr > z =  0.443
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 347.61  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  = 114.37  Prob > chi2 =  0.272
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)   =  75.31  Prob > chi2 =  0.227
    Difference (null H = exogenous): chi2(39)   =  39.06  Prob > chi2 =  0.467
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)   =  42.30  Prob > chi2 =  0.023
    Difference (null H = exogenous): chi2(80)   =  72.07  Prob > chi2 =  0.724
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)   =  97.06  Prob > chi2 =  0.052
    Difference (null H = exogenous): chi2(30)   =  17.31  Prob > chi2 =  0.969
  iv(ea _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(94)   = 103.90  Prob > chi2 =  0.228
    Difference (null H = exogenous): chi2(12)   =  10.47  Prob > chi2 =  0.575

/*The HCASM with SSA and EA dummies*/
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
```

```
Number of instruments = 124                   Obs per group: min =         3
F(17, 133)    =   1401.92                                    avg =      8.73
Prob > F      =     0.000                                    max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9528508   .0161521    58.99   0.000     .9209027     .984799
             |
        ln_s |   .0761813   .0276202     2.76   0.007     .0215497    .1308129
      ln_ngd |  -.1210238   .0756872    -1.60   0.112    -.2707301    .0286825
        ln_h |   .0353268   .0324488     1.09   0.278    -.0288557    .0995093
         ssa |  -.0567555   .0256735    -2.21   0.029    -.1075368   -.0059743
          ea |   .0449592    .024996     1.80   0.074    -.0044818    .0944003
  _Iperiod_2 |   .0839192    .031492     2.66   0.009     .0216293    .1462092
  _Iperiod_3 |   .0467064   .0283078     1.65   0.101    -.0092853    .1026982
  _Iperiod_4 |   .0972525   .0259605     3.75   0.000     .0459036    .1486015
  _Iperiod_5 |   .1056953   .0281712     3.75   0.000     .0499737     .161417
  _Iperiod_6 |   .0405524   .0216941     1.87   0.064    -.0023577    .0834625
  _Iperiod_7 |   .0292567   .0208935     1.40   0.164    -.0120699    .0705833
  _Iperiod_8 |  -.0715881   .0207578    -3.45   0.001    -.1126463   -.0305299
  _Iperiod_9 |  -.0063304   .0205824    -0.31   0.759    -.0470415    .0343807
 _Iperiod_10 |  -.0721553   .0269139    -2.68   0.008      -.12539   -.0189206
 _Iperiod_11 |    .014033    .018759     0.75   0.456    -.0230715    .0511375
 _Iperiod_12 |    .021175   .0130222     1.63   0.106    -.0045825    .0469325
       _cons |   .2732402   .2268011     1.20   0.230    -.1753635     .721844
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ssa ea _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.51  Pr > z =  0.012
Arellano-Bond test for AR(2) in first differences: z =   0.77  Pr > z =  0.442
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 348.79  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  = 114.84  Prob > chi2 =  0.262
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)   =  75.83  Prob > chi2 =  0.215
    Difference (null H = exogenous): chi2(39)   =  39.01  Prob > chi2 =  0.469
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)   =  42.24  Prob > chi2 =  0.023
    Difference (null H = exogenous): chi2(80)   =  72.60  Prob > chi2 =  0.709
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)   =  94.88  Prob > chi2 =  0.070
    Difference (null H = exogenous): chi2(30)   =  19.96  Prob > chi2 =  0.918
  iv(ssa ea _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(93)   = 104.98  Prob > chi2 =  0.186
    Difference (null H = exogenous): chi2(13)   =   9.87  Prob > chi2 =  0.705
```

Appendix II.9: STATA output – SGMM2R_IS6 estimation of the HCASM with SSA and EA dummies – Sample 2

```
/*The HCASM with SSA dummy*/
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1114
Time variable : period                          Number of groups   =       127
Number of instruments = 123                     Obs per group: min =         3
F(16, 126)    =   1204.59                                      avg =      8.77
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9717483   .0145916    66.60   0.000      .942872    1.000625
             |
        ln_s |   .0722974   .0285167     2.54   0.012     .0158636    .1287312
      ln_ngd |   .0509123   .0544627     0.93   0.352    -.0568678    .1586923
        ln_h |   .0437508    .033331     1.31   0.192    -.0222103    .1097119
         ssa |  -.0687028   .0232469    -2.96   0.004    -.1147076    -.022698
  _Iperiod_2 |   .1042904   .0300125     3.47   0.001     .0448965    .1636843
  _Iperiod_3 |   .0608822   .0264989     2.30   0.023     .0084416    .1133228
  _Iperiod_4 |   .1027609   .0255387     4.02   0.000     .0522205    .1533012
  _Iperiod_5 |   .1144764   .0299011     3.83   0.000      .055303    .1736499
  _Iperiod_6 |   .0306743   .0223206     1.37   0.172    -.0134975    .0748461
  _Iperiod_7 |    .019259   .0231228     0.83   0.406    -.0265002    .0650183
  _Iperiod_8 |  -.0577234   .0204781    -2.82   0.006    -.0982489   -.0171979
  _Iperiod_9 |  -.0230155   .0194987    -1.18   0.240     -.061603    .0155719
 _Iperiod_10 |  -.0752311   .0267727    -2.81   0.006    -.1282135   -.0222488
 _Iperiod_11 |   .0029552   .0168904     0.17   0.861    -.0304703    .0363808
 _Iperiod_12 |   .0246371   .0138561     1.78   0.078    -.0027838    .0520579
       _cons |   .5519056   .1809798     3.05   0.003     .1937519    .9100593
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ssa _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.39  Pr > z =  0.017
Arellano-Bond test for AR(2) in first differences: z =  -0.78  Pr > z =  0.434
------------------------------------------------------------------------------
```

```
Sargan test of overid. restrictions: chi2(106)  = 315.59  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  = 117.66  Prob > chi2 =  0.207
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)   =  75.80  Prob > chi2 =  0.216
    Difference (null H = exogenous): chi2(39)   =  41.86  Prob > chi2 =  0.348
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)   =  42.20  Prob > chi2 =  0.023
    Difference (null H = exogenous): chi2(80)   =  75.46  Prob > chi2 =  0.623
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)   =  98.08  Prob > chi2 =  0.045
    Difference (null H = exogenous): chi2(30)   =  19.58  Prob > chi2 =  0.927
  iv(ssa _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(94)   = 107.31  Prob > chi2 =  0.164
    Difference (null H = exogenous): chi2(12)   =  10.35  Prob > chi2 =  0.585

/*The HCASM with EA dummy*/
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period         _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1114
Time variable : period                          Number of groups   =       127
Number of instruments = 123                     Obs per group: min =         3
F(16, 126)    =   1342.60                                      avg =      8.77
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9868521   .0132595    74.43   0.000     .9606119    1.013092
             |
        ln_s |   .0676092   .0275091     2.46   0.015     .0131695    .1220489
      ln_ngd |   .0225654   .0525235     0.43   0.668     -.081377    .1265078
        ln_h |   .0466407    .034919     1.34   0.184     -.022463    .1157443
          ea |   .0514244   .0214913     2.39   0.018     .0088939     .093955
  _Iperiod_2 |   .1246037   .0319655     3.90   0.000     .0613449    .1878625
  _Iperiod_3 |    .081403   .0261391     3.11   0.002     .0296745    .1331316
  _Iperiod_4 |   .1178135   .0258386     4.56   0.000     .0666797    .1689474
  _Iperiod_5 |   .1266375   .0298628     4.24   0.000     .0675399    .1857351
  _Iperiod_6 |    .042677    .022192     1.92   0.057    -.0012404    .0865943
  _Iperiod_7 |    .032459   .0225166     1.44   0.152    -.0121008    .0770187
  _Iperiod_8 |  -.0532998   .0197573    -2.70   0.008    -.0923989   -.0142007
  _Iperiod_9 |  -.0179529   .0191372    -0.94   0.350     -.055825    .0199191
 _Iperiod_10 |  -.0723104   .0248493    -2.91   0.004    -.1214864   -.0231344
 _Iperiod_11 |   .0059824   .0160541     0.37   0.710    -.0257882    .0377529
 _Iperiod_12 |   .0251703   .0136763     1.84   0.068    -.0018948    .0522353
       _cons |   .2930731   .1561005     1.88   0.063    -.0158452    .6019915
------------------------------------------------------------------------------
```

```
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ea _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.30  Pr > z =  0.021
Arellano-Bond test for AR(2) in first differences: z =  -0.79  Pr > z =  0.430
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 323.13  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  = 116.31  Prob > chi2 =  0.232
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)   =  77.43  Prob > chi2 =  0.180
    Difference (null H = exogenous): chi2(39)   =  38.88  Prob > chi2 =  0.475
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)   =  43.22  Prob > chi2 =  0.018
    Difference (null H = exogenous): chi2(80)   =  73.09  Prob > chi2 =  0.695
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)   = 100.46  Prob > chi2 =  0.032
    Difference (null H = exogenous): chi2(30)   =  15.84  Prob > chi2 =  0.984
  iv(ea _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(94)   = 108.92  Prob > chi2 =  0.139
    Difference (null H = exogenous): chi2(12)   =   7.39  Prob > chi2 =  0.831

/*The HCASM with SSA and EA dummies*/
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1114
Time variable : period                          Number of groups   =       127
Number of instruments = 124                     Obs per group: min =         3
F(17, 126)    =   1270.98                                      avg =      8.77
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |    .972943   .0146863    66.25   0.000     .9438792    1.002007
             |
        ln_s |   .0682447   .0278895     2.45   0.016     .0130523    .1234371
```

```
      ln_ngd |   .0442322   .0586742     0.75   0.452    -.0718822    .1603466
        ln_h |     .04451   .0328336     1.36   0.178    -.0204667    .1094867
         ssa |  -.0596411   .0248332    -2.40   0.018    -.1087854   -.0104969
          ea |    .030376   .0223703     1.36   0.177    -.0138943    .0746462
  _Iperiod_2 |   .1084335    .029345     3.70   0.000     .0503606    .1665063
  _Iperiod_3 |   .0643208   .0259142     2.48   0.014     .0130374    .1156043
  _Iperiod_4 |   .1036042   .0254833     4.07   0.000     .0531736    .1540349
  _Iperiod_5 |   .1156804   .0291652     3.97   0.000     .0579634    .1733974
  _Iperiod_6 |   .0324907   .0222038     1.46   0.146    -.0114501    .0764314
  _Iperiod_7 |   .0221485    .022874     0.97   0.335    -.0231185    .0674155
  _Iperiod_8 |  -.0560744   .0198646    -2.82   0.006    -.0953858   -.0167629
  _Iperiod_9 |  -.0222966   .0194649    -1.15   0.254    -.0608171    .0162239
 _Iperiod_10 |  -.0716756   .0256419    -2.80   0.006    -.1224201    -.020931
 _Iperiod_11 |   .0038915     .01636     0.24   0.812    -.0284845    .0362674
 _Iperiod_12 |   .0243638   .0135341     1.80   0.074    -.0024197    .0511473
       _cons |   .5069721   .1974733     2.57   0.011     .1161783     .897766
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ssa ea _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.35  Pr > z =  0.019
Arellano-Bond test for AR(2) in first differences: z =  -0.78  Pr > z =  0.436
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 315.77  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  = 115.41  Prob > chi2 =  0.250
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)   =  76.35  Prob > chi2 =  0.203
    Difference (null H = exogenous): chi2(39)   =  39.06  Prob > chi2 =  0.467
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)   =  43.25  Prob > chi2 =  0.018
    Difference (null H = exogenous): chi2(80)   =  72.16  Prob > chi2 =  0.722
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)   =  98.33  Prob > chi2 =  0.043
    Difference (null H = exogenous): chi2(30)   =  17.08  Prob > chi2 =  0.972
   iv(ssa ea _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(93)   = 106.32  Prob > chi2 =  0.163
    Difference (null H = exogenous): chi2(13)   =   9.10  Prob > chi2 =  0.766
```

Appendix II.10: STATA output – Two-step procedure using SGMM2R_IS3 and SGMM2R_IS6 - Sample 1

```
// Two-step procedure using SGMM2R_IS3
/*Step1*/
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 134                     Obs per group: min =         3
F(15, 133)    =   1039.51                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9491583   .0169551    55.98   0.000     .9156218    .9826949
             |
        ln_s |   .1108652   .0302577     3.66   0.000     .0510166    .1707139
      ln_ngd |  -.2072185   .0623745    -3.32   0.001    -.3305929   -.0838441
        ln_h |    .033007   .0285858     1.15   0.250    -.0235345    .0895486
  _Iperiod_2 |   .0711818   .0274109     2.60   0.010     .0169641    .1253996
  _Iperiod_3 |   .0386761   .0225293     1.72   0.088     -.005886    .0832382
  _Iperiod_4 |   .0961565   .0208097     4.62   0.000     .0549957    .1373172
  _Iperiod_5 |   .1039037    .023427     4.44   0.000     .0575661    .1502413
  _Iperiod_6 |   .0458461   .0226728     2.02   0.045     .0010003     .090692
  _Iperiod_7 |   .0403738   .0222737     1.81   0.072    -.0036828    .0844303
  _Iperiod_8 |  -.0751275   .0180514    -4.16   0.000    -.1108326   -.0394225
  _Iperiod_9 |  -.0025593   .0173546    -0.15   0.883     -.036886    .0317674
 _Iperiod_10 |  -.0887334    .026467    -3.35   0.001    -.1410841   -.0363826
 _Iperiod_11 |   .0140334   .0167751     0.84   0.404    -.0191471    .0472139
 _Iperiod_12 |   .0181871   .0149029     1.22   0.224    -.0112902    .0476644
       _cons |   .1318925   .1570283     0.84   0.402    -.1787034    .4424883
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(1/2).ln_h
    L(2/3).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    D.ln_h
    DL.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.55  Pr > z =  0.011
Arellano-Bond test for AR(2) in first differences: z =   0.77  Pr > z =  0.442
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(118)  = 435.57  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(118)  = 126.91  Prob > chi2 =  0.271
```

```
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(76)   =  94.00  Prob > chi2 =  0.079
    Difference (null H = exogenous): chi2(42)   =  32.91  Prob > chi2 =  0.841
  gmm(ln_y ln_s ln_ngd, lag(2 3))
    Hansen test excluding group:     chi2(28)   =  40.36  Prob > chi2 =  0.061
    Difference (null H = exogenous): chi2(90)   =  86.55  Prob > chi2 =  0.583
  gmm(ln_h, lag(1 2))
    Hansen test excluding group:     chi2(86)   = 104.85  Prob > chi2 =  0.082
    Difference (null H = exogenous): chi2(32)   =  22.05  Prob > chi2 =  0.906
  iv(_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(107)  = 115.10  Prob > chi2 =  0.279
    Difference (null H = exogenous): chi2(11)   =  11.81  Prob > chi2 =  0.378

/*Step2*/
*Regress the residuals on SSA
Linear regression                                      Number of obs =    1170
                                                       F(  1,   133) =    7.20
                                                       Prob > F      =  0.0082
                                                       R-squared     =  0.0149
                                                       Root MSE      =  .16733
                                   (Std. Err. adjusted for 134 clusters in id)
------------------------------------------------------------------------------
             |               Robust
   RES_SGMM2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         ssa |  -.0499723   .0186242    -2.68   0.008    -.0868102   -.0131344
       _cons |   .0104757   .0064228     1.63   0.105    -.0022283    .0231797
------------------------------------------------------------------------------

*Regress the residuals on EA
Linear regression                                      Number of obs =    1170
                                                       F(  1,   133) =   10.67
                                                       Prob > F      =  0.0014
                                                       R-squared     =  0.0133
                                                       Root MSE      =  .16747
                                   (Std. Err. adjusted for 134 clusters in id)
------------------------------------------------------------------------------
             |               Robust
   RES_SGMM2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ea |   .0604469   .0185086     3.27   0.001     .0238375    .0970563
       _cons |  -.0074082    .006857    -1.08   0.282    -.0209712    .0061547
------------------------------------------------------------------------------

*Regress the residuals on SSA and EA
Linear regression                                      Number of obs =    1170
                                                       F(  2,   133) =    7.17
                                                       Prob > F      =  0.0011
                                                       R-squared     =  0.0237
                                                       Root MSE      =  .16665
                                   (Std. Err. adjusted for 134 clusters in id)
------------------------------------------------------------------------------
             |               Robust
   RES_SGMM2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         ssa |  -.0424965    .018692    -2.27   0.025    -.0794684   -.0055245
          ea |   .0500388   .0184207     2.72   0.007     .0136033    .0864743
       _cons |   .0029999   .0065969     0.45   0.650    -.0100486    .0160484
------------------------------------------------------------------------------
```

```
// Two-step procedure using SGMM2R_IS6
/*Step1*/
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 122                     Obs per group: min =         3
F(15, 133)    =   1290.55                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9693327   .0153923    62.98   0.000     .9388873    .9997781
             |
        ln_s |   .0878797   .0283445     3.10   0.002     .0318153     .143944
      ln_ngd |  -.1341353   .0707267    -1.90   0.060    -.2740299    .0057594
        ln_h |   .0295511   .0348958     0.85   0.399    -.0394714    .0985737
  _Iperiod_2 |   .0968734   .0318583     3.04   0.003     .0338589    .1598879
  _Iperiod_3 |   .0571315   .0295933     1.93   0.056    -.0014029    .1156658
  _Iperiod_4 |   .1116336   .0264772     4.22   0.000     .0592626    .1640045
  _Iperiod_5 |   .1188615   .0279948     4.25   0.000     .0634888    .1742342
  _Iperiod_6 |   .0499449   .0212506     2.35   0.020      .007912    .0919778
  _Iperiod_7 |   .0366961   .0208173     1.76   0.080    -.0044796    .0778719
  _Iperiod_8 |  -.0676758    .020985    -3.22   0.002    -.1091832   -.0261683
  _Iperiod_9 |   .0007472   .0192598     0.04   0.969    -.0373479    .0388423
 _Iperiod_10 |  -.0662208    .025535    -2.59   0.011     -.116728   -.0157136
 _Iperiod_11 |   .0199016   .0184698     1.08   0.283     -.016631    .0564342
 _Iperiod_12 |   .0238217   .0134517     1.77   0.079    -.0027852    .0504287
       _cons |   .0974066   .1690253     0.58   0.565    -.2369189    .4317322
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.51  Pr > z =  0.012
Arellano-Bond test for AR(2) in first differences: z =   0.78  Pr > z =  0.438
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 350.80  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  = 114.27  Prob > chi2 =  0.274
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)   =  75.55  Prob > chi2 =  0.222
```

```
    Difference (null H = exogenous): chi2(39)   =  38.72  Prob > chi2 =  0.482
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)   =  41.26  Prob > chi2 =  0.029
    Difference (null H = exogenous): chi2(80)   =  73.01  Prob > chi2 =  0.697
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)   =  95.52  Prob > chi2 =  0.065
    Difference (null H = exogenous): chi2(30)   =  18.76  Prob > chi2 =  0.945
  iv(_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(95)   = 103.58  Prob > chi2 =  0.257
    Difference (null H = exogenous): chi2(11)   =  10.70  Prob > chi2 =  0.469

/*Step2*/
*Regress the residuals on SSA
Linear regression                                      Number of obs =    1170
                                                       F(  1,   133) =    5.52
                                                       Prob > F      =  0.0203
                                                       R-squared     =  0.0099
                                                       Root MSE      =  .16434
                                   (Std. Err. adjusted for 134 clusters in id)
------------------------------------------------------------------------------
             |               Robust
   RES_SGMM2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         ssa |  -.0398243   .0169563    -2.35   0.020    -.0733632   -.0062855
       _cons |   .0076314   .0062351     1.22   0.223    -.0047014    .0199642
------------------------------------------------------------------------------

*Regress the residuals on EA
Linear regression                                      Number of obs =    1170
                                                       F(  1,   133) =   13.10
                                                       Prob > F      =  0.0004
                                                       R-squared     =  0.0158
                                                       Root MSE      =  .16384
                                   (Std. Err. adjusted for 134 clusters in id)
------------------------------------------------------------------------------
             |               Robust
   RES_SGMM2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ea |   .0645497   .0178358     3.62   0.000     .0292711    .0998283
       _cons |  -.0085386   .0062604    -1.36   0.175    -.0209214    .0038443
------------------------------------------------------------------------------

*Regress the residuals on SSA and EA
Linear regression                                      Number of obs =    1170
                                                       F(  2,   133) =    7.67
                                                       Prob > F      =  0.0007
                                                       R-squared     =  0.0217
                                                       Root MSE      =  .16342
                                   (Std. Err. adjusted for 134 clusters in id)
------------------------------------------------------------------------------
             |               Robust
   RES_SGMM2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         ssa |  -.0313268    .016965    -1.85   0.067    -.0648829    .0022292
          ea |   .0568772   .0178359     3.19   0.002     .0215984     .092156
       _cons |  -.0008661   .0062416    -0.14   0.890    -.0132118    .0114796
------------------------------------------------------------------------------
```

Appendix II.11: STATA output – Two-step procedure using SGMM2R_IS6 - Sample 2

```
//Step1
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1114
Time variable : period                          Number of groups   =       127
Number of instruments = 122                     Obs per group: min =         3
F(15, 126)    =   1332.85                                      avg =      8.77
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9875377   .0143357    68.89   0.000     .9591677    1.015908
             |
        ln_s |    .075429   .0286363     2.63   0.009     .0187586    .1320993
      ln_ngd |   .0263655   .0501575     0.53   0.600    -.0728947    .1256257
        ln_h |    .042071    .036585     1.15   0.252    -.0303297    .1144716
  _Iperiod_2 |   .1178152   .0321766     3.66   0.000     .0541387    .1814916
  _Iperiod_3 |   .0763122   .0272835     2.80   0.006      .022319    .1303054
  _Iperiod_4 |   .1160123   .0263936     4.40   0.000     .0637802    .1682443
  _Iperiod_5 |   .1247738   .0313029     3.99   0.000     .0628263    .1867212
  _Iperiod_6 |   .0411138   .0224752     1.83   0.070     -.003364    .0855915
  _Iperiod_7 |   .0296694   .0234618     1.26   0.208    -.0167609    .0760996
  _Iperiod_8 |  -.0545706   .0206072    -2.65   0.009    -.0953517   -.0137896
  _Iperiod_9 |  -.0173998    .019459    -0.89   0.373    -.0559086    .0211089
 _Iperiod_10 |  -.0736548   .0257082    -2.87   0.005    -.1245305   -.0227791
 _Iperiod_11 |   .0059251   .0166139     0.36   0.722    -.0269534    .0388036
 _Iperiod_12 |   .0254468   .0139933     1.82   0.071    -.0022456    .0531392
       _cons |   .3258396   .1587474     2.05   0.042     .0116832    .6399961
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.34  Pr > z =  0.020
Arellano-Bond test for AR(2) in first differences: z =  -0.79  Pr > z =  0.430
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 324.28  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  = 117.75  Prob > chi2 =  0.205
  (Robust, but weakened by many instruments.)
```

```
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)    =  76.66  Prob > chi2 =  0.196
    Difference (null H = exogenous): chi2(39)    =  41.09  Prob > chi2 =  0.379
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)    =  41.47  Prob > chi2 =  0.028
    Difference (null H = exogenous): chi2(80)    =  76.27  Prob > chi2 =  0.597
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)    =  99.71  Prob > chi2 =  0.035
    Difference (null H = exogenous): chi2(30)    =  18.04  Prob > chi2 =  0.958
  iv(_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(95)    = 109.94  Prob > chi2 =  0.140
    Difference (null H = exogenous): chi2(11)    =   7.81  Prob > chi2 =  0.730

//Step2
/*Regress the residuals on SSA*/
Linear regression                                      Number of obs =    1114
                                                       F(  1,   126) =    4.60
                                                       Prob > F      =  0.0338
                                                       R-squared     =  0.0097
                                                       Root MSE      =  .15658
                                  (Std. Err. adjusted for 127 clusters in id)
------------------------------------------------------------------------------
             |               Robust
   RES_SGMM2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         ssa |   -.037368   .0174183    -2.15   0.034    -.0718382   -.0028977
       _cons |   .0063079   .0066131     0.95   0.342    -.0067793    .0193951
------------------------------------------------------------------------------

/*Regress the residuals on EA*/
Linear regression                                      Number of obs =    1114
                                                       F(  1,   126) =    5.86
                                                       Prob > F      =  0.0169
                                                       R-squared     =  0.0100
                                                       Root MSE      =  .15656
                                  (Std. Err. adjusted for 127 clusters in id)
------------------------------------------------------------------------------
             |               Robust
   RES_SGMM2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ea |   .0478912   .0197825     2.42   0.017     .0087422    .0870402
       _cons |  -.0078671    .006583    -1.20   0.234    -.0208947    .0051604
------------------------------------------------------------------------------

/*Regress the residuals on SSA and EA*/
Linear regression                                      Number of obs =    1114
                                                       F(  2,   126) =    4.16
                                                       Prob > F      =  0.0179
                                                       R-squared     =  0.0165
                                                       Root MSE      =  .15612
                                  (Std. Err. adjusted for 127 clusters in id)
------------------------------------------------------------------------------
             |               Robust
   RES_SGMM2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         ssa |  -.0310404   .0174629    -1.78   0.078    -.0655989    .0035181
          ea |   .0400437   .0198338     2.02   0.046     .0007932    .0792942
       _cons |  -.0000196   .0067123    -0.00   0.998     -.013303    .0132638
------------------------------------------------------------------------------
```

Appendix II.12: STATA output – SGMM2R_IS3 and SGMM2R_IS6 estimation of the HCASM with SSA and EA dummies and interactions SSAP and EAP – Sample 1

```
//SGMM2R_IS3
/*The HCASM with SSA and SSAP*/
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
ssap2 dropped due to collinearity
ssap3 dropped due to collinearity
ssap13 dropped due to collinearity
_Iperiod_13 dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 144                     Obs per group: min =         3
F(25, 133)    =   1316.25                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9203842   .0193751    47.50   0.000     .8820611    .9587074
             |
        ln_s |   .0942471   .0320427     2.94   0.004     .0308678    .1576264
      ln_ngd |   -.141728   .0701471    -2.02   0.045    -.2804762   -.0029798
        ln_h |    .050607   .0290175     1.74   0.083    -.0067885    .1080024
         ssa |  -.0724223   .0327649    -2.21   0.029      -.13723   -.0076146
       ssap4 |  -.0172718   .0466291    -0.37   0.712    -.1095024    .0749588
       ssap5 |   -.040176   .0451688    -0.89   0.375    -.1295182    .0491662
       ssap6 |  -.0338269   .0470559    -0.72   0.473    -.1269016    .0592478
       ssap7 |  -.0692937   .0523885    -1.32   0.188     -.172916    .0343287
       ssap8 |   .0386924   .0471084     0.82   0.413    -.0544861     .131871
       ssap9 |  -.0062479   .0420948    -0.15   0.882    -.0895097     .077014
      ssap10 |  -.0488742   .0494005    -0.99   0.324    -.1465864     .048838
      ssap11 |  -.0465574    .046795    -0.99   0.322    -.1391161    .0460013
      ssap12 |  -.0589663   .0296639    -1.99   0.049    -.1176405   -.0002922
  _Iperiod_2 |   .0629039   .0278317     2.26   0.025     .0078538    .1179539
  _Iperiod_3 |   .0368681   .0262407     1.40   0.162     -.015035    .0887712
  _Iperiod_4 |   .0856794    .024411     3.51   0.001     .0373954    .1339634
  _Iperiod_5 |   .1051099   .0248268     4.23   0.000     .0560034    .1542164
  _Iperiod_6 |   .0501852   .0213537     2.35   0.020     .0079484     .092422
  _Iperiod_7 |   .0527291   .0219285     2.40   0.018     .0093554    .0961028
  _Iperiod_8 |  -.0810302   .0192886    -4.20   0.000    -.1191824    -.042878
  _Iperiod_9 |   -.008836   .0186421    -0.47   0.636    -.0457093    .0280373
 _Iperiod_10 |  -.0699721   .0281732    -2.48   0.014    -.1256977   -.0142465
 _Iperiod_11 |     .01908   .0161161     1.18   0.239    -.0127969     .050957
 _Iperiod_12 |   .0340065   .0173013     1.97   0.051    -.0002149    .0682278
       _cons |   .5423868   .2257468     2.40   0.018     .0958683    .9889052
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
```

```
    L(1/2).ln_h
    L(2/3).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ssa ssap2 ssap3 ssap4 ssap5 ssap6 ssap7 ssap8 ssap9 ssap10
    ssap11 ssap12 ssap13 _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5
    _Iperiod_6 _Iperiod_7 _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11
    _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    D.ln_h
    DL.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.53  Pr > z =  0.012
Arellano-Bond test for AR(2) in first differences: z =   0.87  Pr > z =  0.385
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(118)  = 449.48  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(118)  = 117.74  Prob > chi2 =  0.489
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(76)   =  94.65  Prob > chi2 =  0.073
    Difference (null H = exogenous): chi2(42)   =  23.10  Prob > chi2 =  0.992
  gmm(ln_y ln_s ln_ngd, lag(2 3))
    Hansen test excluding group:     chi2(28)   =  29.04  Prob > chi2 =  0.411
    Difference (null H = exogenous): chi2(90)   =  88.70  Prob > chi2 =  0.519
  gmm(ln_h, lag(1 2))
    Hansen test excluding group:     chi2(86)   = 110.00  Prob > chi2 =  0.042
    Difference (null H = exogenous): chi2(32)   =   7.75  Prob > chi2 =  1.000
  iv(ssa ssap2 ssap3 ssap4 ssap5 ssap6 ssap7 ssap8 ssap9 ssap10 ssap11 ssap12
ssap13 _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(97)   = 109.39  Prob > chi2 =  0.184
    Difference (null H = exogenous): chi2(21)   =   8.36  Prob > chi2 =  0.993

/*The HCASM with EA and EAP*/
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
eap2 dropped due to collinearity
eap13 dropped due to collinearity
_Iperiod_13 dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 145                     Obs per group: min =         3
F(26, 133)    =   1062.38                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
```

```
        ln_y |
         L1. |   .9571003   .0161086    59.42   0.000     .9252382    .9889625
             |
        ln_s |   .0943832     .02717     3.47   0.001     .0406421    .1481243
      ln_ngd |  -.2365492   .0547604    -4.32   0.000    -.3448631   -.1282353
        ln_h |   .0171147   .0293038     0.58   0.560    -.0408471    .0750764
          ea |   .0881995    .032844     2.69   0.008     .0232353    .1531637
        eap3 |  -.0233533   .0918772    -0.25   0.800    -.2050828    .1583763
        eap4 |  -.1278697   .0618561    -2.07   0.041    -.2502186   -.0055208
        eap5 |  -.0396525   .0678139    -0.58   0.560    -.1737857    .0944807
        eap6 |  -.0636569   .0831669    -0.77   0.445    -.2281577    .1008439
        eap7 |  -.0288927   .0766445    -0.38   0.707    -.1804924    .1227071
        eap8 |   .0497237   .0409802     1.21   0.227    -.0313336     .130781
        eap9 |   .0018238   .0463615     0.04   0.969    -.0898775    .0935251
       eap10 |   .0988027   .0624198     1.58   0.116    -.0246612    .2222666
       eap11 |  -.1095805   .0436375    -2.51   0.013    -.1958937   -.0232673
       eap12 |   .0033402   .0169705     0.20   0.844    -.0302269    .0369073
  _Iperiod_2 |    .069615   .0268268     2.59   0.011     .0165526    .1226773
  _Iperiod_3 |   .0383126   .0251943     1.52   0.131    -.0115207     .088146
  _Iperiod_4 |   .0995105   .0231218     4.30   0.000     .0537764    .1452445
  _Iperiod_5 |   .0959182   .0222089     4.32   0.000     .0519899    .1398465
  _Iperiod_6 |   .0467972    .024293     1.93   0.056    -.0012535    .0948478
  _Iperiod_7 |   .0438653   .0220042     1.99   0.048     .0003419    .0873887
  _Iperiod_8 |  -.0868897   .0195725    -4.44   0.000    -.1256035    -.048176
  _Iperiod_9 |  -.0101738   .0166216    -0.61   0.542    -.0430507    .0227032
 _Iperiod_10 |  -.0968527   .0262262    -3.69   0.000     -.148727   -.0449784
 _Iperiod_11 |   .0231263   .0175884     1.31   0.191    -.0116629    .0579154
 _Iperiod_12 |   .0159447   .0146902     1.09   0.280    -.0131119    .0450013
       _cons |  -.0296421   .1635014    -0.18   0.856    -.3530414    .2937573
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(1/2).ln_h
    L(2/3).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ea eap2 eap3 eap4 eap5 eap6 eap7 eap8 eap9 eap10 eap11 eap12
    eap13 _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    D.ln_h
    DL.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.55  Pr > z =  0.011
Arellano-Bond test for AR(2) in first differences: z =   0.71  Pr > z =  0.481
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(118)  = 417.69  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(118)  = 112.18  Prob > chi2 =  0.634
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(76)   =  87.40  Prob > chi2 =  0.175
    Difference (null H = exogenous): chi2(42)   =  24.78  Prob > chi2 =  0.984
  gmm(ln_y ln_s ln_ngd, lag(2 3))
    Hansen test excluding group:     chi2(28)   =  41.60  Prob > chi2 =  0.047
    Difference (null H = exogenous): chi2(90)   =  70.58  Prob > chi2 =  0.935
  gmm(ln_h, lag(1 2))
    Hansen test excluding group:     chi2(86)   = 102.70  Prob > chi2 =  0.106
    Difference (null H = exogenous): chi2(32)   =   9.47  Prob > chi2 =  1.000
```

```
  iv(ea eap2 eap3 eap4 eap5 eap6 eap7 eap8 eap9 eap10 eap11 eap12 eap13
_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7 _Iperiod_8
_Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13, eq(level))
    Hansen test excluding group:     chi2(96)   = 107.28  Prob > chi2 =  0.203
    Difference (null H = exogenous): chi2(22)   =   4.90  Prob > chi2 =  1.000

/*The HCASM with SSA, EA, SSAP and EAP*/
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
ssap2 dropped due to collinearity
ssap3 dropped due to collinearity
ssap13 dropped due to collinearity
eap2 dropped due to collinearity
eap13 dropped due to collinearity
_Iperiod_13 dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 155                     Obs per group: min =         3
F(36, 133)    =   1141.64                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9440486   .0230319    40.99   0.000     .8984924    .9896048
             |
        ln_s |   .0758957   .0281629     2.69   0.008     .0201905    .1316009
      ln_ngd |  -.1902365   .0640044    -2.97   0.004    -.3168348   -.0636382
        ln_h |   .0260675   .0350114     0.74   0.458    -.0431837    .0953186
         ssa |  -.0194043   .0346092    -0.56   0.576      -.08786    .0490514
          ea |   .1174195   .0659257     1.78   0.077    -.0129789    .2478179
       ssap4 |  -.0559741   .0496891    -1.13   0.262    -.1542571    .0423089
       ssap5 |  -.0707453   .0462381    -1.53   0.128    -.1622025    .0207119
       ssap6 |  -.0666102   .0466415    -1.43   0.156    -.1588652    .0256449
       ssap7 |  -.1119356   .0501735    -2.23   0.027    -.2111768   -.0126943
       ssap8 |   .0050054   .0499975     0.10   0.920    -.0938877    .1038986
       ssap9 |  -.0403042   .0434668    -0.93   0.355    -.1262798    .0456715
      ssap10 |  -.0418139    .051086    -0.82   0.415      -.14286    .0592323
      ssap11 |  -.0837606   .0430067    -1.95   0.054    -.1688261    .0013049
      ssap12 |  -.0733667   .0347575    -2.11   0.037    -.1421156   -.0046178
        eap3 |  -.0016996   .1779542    -0.01   0.992    -.3536861    .3502868
        eap4 |   -.183798   .0805016    -2.28   0.024     -.343027    -.024569
        eap5 |  -.0822151   .0861387    -0.95   0.342    -.2525942     .088164
        eap6 |  -.0961377   .0923227    -1.04   0.300    -.2787484     .086473
        eap7 |   -.076001   .0902997    -0.84   0.401    -.2546103    .1026083
        eap8 |   .0085623   .0668337     0.13   0.898    -.1236322    .1407568
        eap9 |  -.0422978   .0703748    -0.60   0.549    -.1814965    .0969009
       eap10 |   .0652984   .0733313     0.89   0.375    -.0797481    .2103448
       eap11 |  -.1467506    .069422    -2.11   0.036    -.2840646   -.0094366
       eap12 |  -.1125601   .1090646    -1.03   0.304    -.3282857    .1031655
  _Iperiod_2 |     .06985   .0259483     2.69   0.008     .0185252    .1211748
```

```
  _Iperiod_3 |   .0429499   .0254609     1.69   0.094    -.0074108    .0933107
  _Iperiod_4 |   .1050738   .0256286     4.10   0.000     .0543813    .1557663
  _Iperiod_5 |   .1126991   .0254437     4.43   0.000     .0623726    .1630257
  _Iperiod_6 |     .05697   .0247165     2.30   0.023     .0080817    .1058583
  _Iperiod_7 |   .0656211    .027699     2.37   0.019     .0108335    .1204086
  _Iperiod_8 |  -.0761402   .0229646    -3.32   0.001    -.1215634    -.030717
  _Iperiod_9 |   .0023272   .0240803     0.10   0.923    -.0453027     .049957
 _Iperiod_10 |  -.0752072    .031657    -2.38   0.019    -.1378235   -.0125909
 _Iperiod_11 |   .0446751   .0192658     2.32   0.022     .0065681     .082782
 _Iperiod_12 |   .0492522   .0239344     2.06   0.042     .0019108    .0965936
       _cons |   .1756458   .2441401     0.72   0.473    -.3072539    .6585455
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(1/2).ln_h
    L(2/3).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ssa ea ssap2 ssap3 ssap4 ssap5 ssap6 ssap7 ssap8 ssap9 ssap10
    ssap11 ssap12 ssap13 eap2 eap3 eap4 eap5 eap6 eap7 eap8 eap9
    eap10 eap11 eap12 eap13 _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5
    _Iperiod_6 _Iperiod_7 _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11
    _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    D.ln_h
    DL.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.51  Pr > z =  0.012
Arellano-Bond test for AR(2) in first differences: z =   0.79  Pr > z =  0.430
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(118)  = 432.00  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(118)  = 100.08  Prob > chi2 =  0.883
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(76)   =  85.89  Prob > chi2 =  0.205
    Difference (null H = exogenous): chi2(42)   =  14.18  Prob > chi2 =  1.000
  gmm(ln_y ln_s ln_ngd, lag(2 3))
    Hansen test excluding group:     chi2(28)   =  31.47  Prob > chi2 =  0.296
    Difference (null H = exogenous): chi2(90)   =  68.60  Prob > chi2 =  0.955
  gmm(ln_h, lag(1 2))
    Hansen test excluding group:     chi2(86)   = 104.64  Prob > chi2 =  0.084
    Difference (null H = exogenous): chi2(32)   =  -4.57  Prob > chi2 =  1.000
  iv(ssa ea ssap2 ssap3 ssap4 ssap5 ssap6 ssap7 ssap8 ssap9 ssap10 ssap11
ssap12 ssap13 eap2 eap3 eap4 eap5 eap6 eap7 eap8 eap9 eap10 eap11 eap12 eap13
_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7 _Iperiod_8
_Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13, eq(level))
    Hansen test excluding group:     chi2(86)   =  96.23  Prob > chi2 =  0.211
    Difference (null H = exogenous): chi2(32)   =   3.85  Prob > chi2 =  1.000
```

```
//SGMM2R_IS6
/*The HCASM with SSA and SSAP*/
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
ssap2 dropped due to collinearity
ssap3 dropped due to collinearity
```

```
ssap13 dropped due to collinearity
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 132                     Obs per group: min =         3
F(25, 133)    =   1728.31                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9403775   .0183194    51.33   0.000     .9041425    .9766125
             |
        ln_s |   .0818259   .0224134     3.65   0.000     .0374931    .1261586
      ln_ngd |  -.0834997   .0675402    -1.24   0.219    -.2170917    .0500922
        ln_h |   .0555984   .0385557     1.44   0.152    -.0206632      .13186
         ssa |  -.0363492   .0330801    -1.10   0.274    -.1017803     .029082
       ssap4 |  -.0247123    .050704    -0.49   0.627    -.1250028    .0755783
       ssap5 |  -.0312887   .0520259    -0.60   0.549    -.1341939    .0716165
       ssap6 |  -.0584573   .0469083    -1.25   0.215    -.1512402    .0343255
       ssap7 |    -.11691   .0509311    -2.30   0.023    -.2176497   -.0161704
       ssap8 |   .0066889   .0453822     0.15   0.883    -.0830753    .0964532
       ssap9 |  -.0116724   .0419211    -0.28   0.781    -.0945908     .071246
      ssap10 |  -.0764421   .0491255    -1.56   0.122    -.1736104    .0207262
      ssap11 |  -.0583116   .0415945    -1.40   0.163     -.140584    .0239608
      ssap12 |  -.0650127   .0310074    -2.10   0.038    -.1263442   -.0036811
  _Iperiod_2 |   .1002566   .0295882     3.39   0.001     .0417323     .158781
  _Iperiod_3 |   .0609134   .0259426     2.35   0.020     .0096001    .1122268
  _Iperiod_4 |   .1126172   .0245724     4.58   0.000     .0640139    .1612204
  _Iperiod_5 |   .1233092   .0281904     4.37   0.000     .0675497    .1790687
  _Iperiod_6 |   .0568619   .0227287     2.50   0.014     .0119054    .1018183
  _Iperiod_7 |   .0488026   .0236011     2.07   0.041     .0021205    .0954847
  _Iperiod_8 |  -.0780589   .0218905    -3.57   0.001    -.1213575   -.0347604
  _Iperiod_9 |  -.0088402   .0206928    -0.43   0.670    -.0497698    .0320895
 _Iperiod_10 |  -.0515753   .0272125    -1.90   0.060    -.1054006    .0022501
 _Iperiod_11 |   .0243368   .0171808     1.42   0.159    -.0096462    .0583198
 _Iperiod_12 |   .0368626   .0155101     2.38   0.019     .0061842    .0675409
       _cons |   .4660217   .2255757     2.07   0.041     .0198416    .9122018
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ssa ssap2 ssap3 ssap4 ssap5 ssap6 ssap7 ssap8 ssap9 ssap10
    ssap11 ssap12 ssap13 _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5
    _Iperiod_6 _Iperiod_7 _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11
    _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.56  Pr > z =  0.011
Arellano-Bond test for AR(2) in first differences: z =   0.87  Pr > z =  0.385
```

```
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 350.77  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  = 113.91  Prob > chi2 =  0.282
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)   =  77.64  Prob > chi2 =  0.176
    Difference (null H = exogenous): chi2(39)   =  36.28  Prob > chi2 =  0.595
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)   =  30.87  Prob > chi2 =  0.233
    Difference (null H = exogenous): chi2(80)   =  83.04  Prob > chi2 =  0.386
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)   =  88.42  Prob > chi2 =  0.156
    Difference (null H = exogenous): chi2(30)   =  25.49  Prob > chi2 =  0.701
  iv(ssa ssap2 ssap3 ssap4 ssap5 ssap6 ssap7 ssap8 ssap9 ssap10 ssap11 ssap12
ssap13 _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(85)   =  97.37  Prob > chi2 =  0.169
    Difference (null H = exogenous): chi2(21)   =  16.54  Prob > chi2 =  0.738

/*The HCASM with EA and EAP*/
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
eap2 dropped due to collinearity
eap13 dropped due to collinearity
_Iperiod_13 dropped due to collinearity
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 133                     Obs per group: min =         3
F(26, 133)    =    887.47                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9717827   .0150612    64.52   0.000     .9419922    1.001573
             |
        ln_s |   .0828566   .0271884     3.05   0.003     .0290791    .1366341
      ln_ngd |  -.1585887   .0658831    -2.41   0.017    -.2889028   -.0282745
        ln_h |   .0153973   .0369181     0.42   0.677    -.0576253      .08842
          ea |   .0755443   .0282438     2.67   0.008     .0196792    .1314094
        eap3 |  -.0822289   .1654119    -0.50   0.620    -.4094071    .2449494
        eap4 |  -.0845725   .0598835    -1.41   0.160    -.2030197    .0338746
        eap5 |  -.0443772   .0880994    -0.50   0.615    -.2186344    .1298799
        eap6 |  -.0450079   .0662512    -0.68   0.498    -.1760501    .0860344
        eap7 |   .0268926   .0659766     0.41   0.684    -.1036065    .1573918
        eap8 |   .0460698   .0470668     0.98   0.329    -.0470266    .1391662
        eap9 |   .0204688    .049567     0.41   0.680    -.0775727    .1185103
       eap10 |   .1189548   .0610096     1.95   0.053    -.0017197    .2396294
       eap11 |  -.0980018    .033363    -2.94   0.004    -.1639926    -.032011
```

```
       eap12 |   .0083604    .025515     0.33   0.744    -.0421072     .058828
  _Iperiod_2 |   .0850859   .0319526     2.66   0.009     .0218848     .148287
  _Iperiod_3 |       .055   .0321888     1.71   0.090    -.0086682    .1186681
  _Iperiod_4 |   .1061222   .0313713     3.38   0.001      .044071    .1681734
  _Iperiod_5 |   .1060304   .0346422     3.06   0.003     .0375094    .1745513
  _Iperiod_6 |   .0463959   .0241237     1.92   0.057    -.0013199    .0941117
  _Iperiod_7 |   .0233487   .0248503     0.94   0.349    -.0258044    .0725017
  _Iperiod_8 |  -.0835545   .0268892    -3.11   0.002    -.1367403   -.0303688
  _Iperiod_9 |  -.0062655   .0214544    -0.29   0.771    -.0487015    .0361706
 _Iperiod_10 |  -.0839688   .0287612    -2.92   0.004    -.1408573   -.0270804
 _Iperiod_11 |   .0216805   .0178901     1.21   0.228    -.0137055    .0570665
 _Iperiod_12 |   .0130359   .0158141     0.82   0.411    -.0182437    .0443155
       _cons |   .0240889   .1729895     0.14   0.889    -.3180776    .3662555
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ea eap2 eap3 eap4 eap5 eap6 eap7 eap8 eap9 eap10 eap11 eap12
    eap13 _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.52  Pr > z =  0.012
Arellano-Bond test for AR(2) in first differences: z =   0.73  Pr > z =  0.464
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 343.47  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  = 104.15  Prob > chi2 =  0.533
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)   =  72.89  Prob > chi2 =  0.291
    Difference (null H = exogenous): chi2(39)   =  31.26  Prob > chi2 =  0.807
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)   =  40.59  Prob > chi2 =  0.034
    Difference (null H = exogenous): chi2(80)   =  63.56  Prob > chi2 =  0.911
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)   =  91.43  Prob > chi2 =  0.109
    Difference (null H = exogenous): chi2(30)   =  12.71  Prob > chi2 =  0.998
  iv(ea eap2 eap3 eap4 eap5 eap6 eap7 eap8 eap9 eap10 eap11 eap12 eap13
_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7 _Iperiod_8
_Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13, eq(level))
    Hansen test excluding group:     chi2(84)   =  91.56  Prob > chi2 =  0.268
    Difference (null H = exogenous): chi2(22)   =  12.59  Prob > chi2 =  0.944


/*The HCASM with SSA, EA, SSAP and EAP*/
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
ssap2 dropped due to collinearity
ssap3 dropped due to collinearity
ssap13 dropped due to collinearity
eap2 dropped due to collinearity
```

```
eap13 dropped due to collinearity
_Iperiod_13 dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1170
Time variable : period                          Number of groups   =       134
Number of instruments = 143                     Obs per group: min =         3
F(36, 133)    =   1479.05                                      avg =      8.73
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9466133   .0235762    40.15   0.000     .8999804    .9932461
             |
        ln_s |   .0706595   .0257381     2.75   0.007     .0197505    .1215685
      ln_ngd |  -.1053355    .075877    -1.39   0.167    -.2554173    .0447463
        ln_h |   .0408753   .0413572     0.99   0.325    -.0409276    .1226781
         ssa |  -.0273851   .0411761    -0.67   0.507    -.1088299    .0540597
          ea |   .0502383   .0340929     1.47   0.143     -.017196    .1176727
       ssap4 |  -.0581002   .0442682    -1.31   0.192    -.1456609    .0294606
       ssap5 |  -.0576165   .0446502    -1.29   0.199    -.1459328    .0306998
       ssap6 |  -.0725336   .0494972    -1.47   0.145    -.1704371    .0253699
       ssap7 |  -.1086304   .0595197    -1.83   0.070    -.2263581    .0090972
       ssap8 |  -.0010083   .0460157    -0.02   0.983    -.0920255     .090009
       ssap9 |  -.0269248   .0432572    -0.62   0.535    -.1124858    .0586362
      ssap10 |  -.0595928   .0512572    -1.16   0.247    -.1609775     .041792
      ssap11 |  -.0835699   .0408053    -2.05   0.043    -.1642812   -.0028586
      ssap12 |  -.0597231   .0344187    -1.74   0.085     -.127802    .0083558
        eap3 |  -.1683767   .1784589    -0.94   0.347    -.5213615    .1846082
        eap4 |  -.1172246   .0589767    -1.99   0.049    -.2338782    -.000571
        eap5 |  -.0556912   .0872085    -0.64   0.524    -.2281862    .1168038
        eap6 |  -.0574316   .0723609    -0.79   0.429    -.2005587    .0856955
        eap7 |  -.0078068   .0668986    -0.12   0.907    -.1401296     .124516
        eap8 |   .0239779   .0546657     0.44   0.662    -.0841488    .1321045
        eap9 |   .0202373   .0510906     0.40   0.693    -.0808181    .1212926
       eap10 |   .1132096    .061466     1.84   0.068    -.0083677     .234787
       eap11 |   -.097782    .033595    -2.91   0.004    -.1642316   -.0313325
       eap12 |  -.0082021   .0327381    -0.25   0.803    -.0729568    .0565526
  _Iperiod_2 |   .0984924   .0328496     3.00   0.003     .0335172    .1634675
  _Iperiod_3 |   .0707669   .0316917     2.23   0.027     .0080821    .1334518
  _Iperiod_4 |   .1241725   .0297928     4.17   0.000     .0652434    .1831016
  _Iperiod_5 |   .1261549   .0335771     3.76   0.000     .0597407    .1925691
  _Iperiod_6 |   .0644529   .0230197     2.80   0.006     .0189208     .109985
  _Iperiod_7 |   .0500033   .0256352     1.95   0.053     -.000702    .1007087
  _Iperiod_8 |  -.0836864   .0289061    -2.90   0.004    -.1408614   -.0265113
  _Iperiod_9 |  -.0119733   .0235942    -0.51   0.613    -.0586418    .0346951
 _Iperiod_10 |  -.0716248   .0304635    -2.35   0.020    -.1318804   -.0113693
 _Iperiod_11 |   .0359093   .0168297     2.13   0.035     .0026209    .0691977
 _Iperiod_12 |   .0313798    .020665     1.52   0.131    -.0094947    .0722543
       _cons |   .3503094   .3133538     1.12   0.266    -.2694924    .9701111
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
```

```
  Standard
    ssa ea ssap2 ssap3 ssap4 ssap5 ssap6 ssap7 ssap8 ssap9 ssap10
    ssap11 ssap12 ssap13 eap2 eap3 eap4 eap5 eap6 eap7 eap8 eap9
    eap10 eap11 eap12 eap13 _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5
    _Iperiod_6 _Iperiod_7 _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11
    _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.53  Pr > z =  0.011
Arellano-Bond test for AR(2) in first differences: z =   0.85  Pr > z =  0.393
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 340.98  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  = 102.75  Prob > chi2 =  0.571
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)   =  79.91  Prob > chi2 =  0.134
    Difference (null H = exogenous): chi2(39)   =  22.83  Prob > chi2 =  0.982
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)   =  30.46  Prob > chi2 =  0.249
    Difference (null H = exogenous): chi2(80)   =  72.29  Prob > chi2 =  0.718
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)   =  87.84  Prob > chi2 =  0.166
    Difference (null H = exogenous): chi2(30)   =  14.90  Prob > chi2 =  0.990
  iv(ssa ea ssap2 ssap3 ssap4 ssap5 ssap6 ssap7 ssap8 ssap9 ssap10 ssap11
ssap12 ssap13 eap2 eap3 eap4 eap5 eap6 eap7 eap8 eap9 eap10 eap11 eap12 eap13
_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7 _Iperiod_8
_Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13, eq(level))
    Hansen test excluding group:     chi2(74)   =  82.93  Prob > chi2 =  0.223
    Difference (null H = exogenous): chi2(32)   =  19.82  Prob > chi2 =  0.954
```

Appendix II.13: STATA output – SGMM2R_IS6 estimation of the HCASM with SSA and EA dummies and interactions SSAP and EAP – Sample 2

```
//The HCASM with SSA and SSAP
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
ssap2 dropped due to collinearity
ssap3 dropped due to collinearity
ssap13 dropped due to collinearity
_Iperiod_13 dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1114
Time variable : period                          Number of groups   =       127
Number of instruments = 132                     Obs per group: min =         3
F(25, 126)    =   1282.12                                      avg =      8.77
```

```
Prob > F      =     0.000                                       max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9647704   .0161069    59.90   0.000     .9328953    .9966455
             |
        ln_s |   .0729427   .0339176     2.15   0.033     .0058208    .1400645
      ln_ngd |   .0646898   .0606806     1.07   0.288    -.0553953     .184775
        ln_h |   .0548874   .0380288     1.44   0.151    -.0203705    .1301452
         ssa |  -.0242439   .0280769    -0.86   0.390    -.0798073    .0313195
       ssap4 |  -.0212314   .0445783    -0.48   0.635    -.1094505    .0669877
       ssap5 |  -.0255802   .0483232    -0.53   0.597    -.1212104    .0700499
       ssap6 |  -.0363158   .0526253    -0.69   0.491    -.1404596     .067828
       ssap7 |  -.0928713    .049257    -1.89   0.062    -.1903494    .0046068
       ssap8 |  -.0321516   .0502402    -0.64   0.523    -.1315755    .0672723
       ssap9 |  -.0184361   .0421618    -0.44   0.663    -.1018731    .0650008
      ssap10 |  -.1137105   .0463099    -2.46   0.015    -.2053564   -.0220647
      ssap11 |  -.0674984   .0479209    -1.41   0.161    -.1623324    .0273356
      ssap12 |  -.0914623   .0332575    -2.75   0.007    -.1572779   -.0256467
  _Iperiod_2 |   .1181104    .028137     4.20   0.000      .062428    .1737927
  _Iperiod_3 |   .0766493   .0253201     3.03   0.003     .0265415    .1267571
  _Iperiod_4 |   .1167409   .0230881     5.06   0.000     .0710502    .1624316
  _Iperiod_5 |   .1326433   .0275837     4.81   0.000      .078056    .1872305
  _Iperiod_6 |   .0473822    .020384     2.32   0.022      .007043    .0877214
  _Iperiod_7 |   .0434107   .0205735     2.11   0.037     .0026964     .084125
  _Iperiod_8 |  -.0479656    .022262    -2.15   0.033    -.0920214   -.0039098
  _Iperiod_9 |  -.0150922   .0201649    -0.75   0.456    -.0549981    .0248136
 _Iperiod_10 |  -.0358413   .0294067    -1.22   0.225    -.0940363    .0223537
 _Iperiod_11 |   .0223891   .0161072     1.39   0.167    -.0094865    .0542646
 _Iperiod_12 |   .0454315   .0165896     2.74   0.007     .0126012    .0782617
       _cons |   .6198432   .2073501     2.99   0.003     .2095035    1.030183
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ssa ssap2 ssap3 ssap4 ssap5 ssap6 ssap7 ssap8 ssap9 ssap10
    ssap11 ssap12 ssap13 _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5
    _Iperiod_6 _Iperiod_7 _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11
    _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.50  Pr > z =  0.013
Arellano-Bond test for AR(2) in first differences: z =  -0.87  Pr > z =  0.384
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 313.78  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  = 103.66  Prob > chi2 =  0.546
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)    =  70.40  Prob > chi2 =  0.365
    Difference (null H = exogenous): chi2(39)    =  33.27  Prob > chi2 =  0.728
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)    =  32.02  Prob > chi2 =  0.192
    Difference (null H = exogenous): chi2(80)    =  71.64  Prob > chi2 =  0.736
```

```
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)   =  88.54  Prob > chi2 =  0.154
    Difference (null H = exogenous): chi2(30)   =  15.12  Prob > chi2 =  0.989
  iv(ssa ssap2 ssap3 ssap4 ssap5 ssap6 ssap7 ssap8 ssap9 ssap10 ssap11 ssap12
ssap13 _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
_Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13,
eq(level))
    Hansen test excluding group:     chi2(85)   =  91.18  Prob > chi2 =  0.304
    Difference (null H = exogenous): chi2(21)   =  12.48  Prob > chi2 =  0.926

//The HCASM with EA and EAP
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
eap2 dropped due to collinearity
eap13 dropped due to collinearity
_Iperiod_13 dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1114
Time variable : period                          Number of groups   =       127
Number of instruments = 133                     Obs per group: min =         3
F(26, 126)    =   1163.47                                      avg =      8.77
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9858461   .0127633    77.24   0.000     .9605878    1.011104
             |
        ln_s |   .0730289   .0273177     2.67   0.009      .018968    .1270897
      ln_ngd |   .0002412    .056059     0.00   0.997    -.1106979    .1111802
        ln_h |    .033973   .0366582     0.93   0.356    -.0385725    .1065184
          ea |   .0639173   .0376797     1.70   0.092    -.0106497    .1384844
        eap3 |  -.0943064   .1924352    -0.49   0.625      -.47513    .2865172
        eap4 |  -.1116439   .0660714    -1.69   0.094    -.2423972    .0191094
        eap5 |  -.0850389   .0979332    -0.87   0.387    -.2788459     .108768
        eap6 |  -.0365049   .0812239    -0.45   0.654    -.1972446    .1242349
        eap7 |   .0125389   .0698221     0.18   0.858     -.125637    .1507149
        eap8 |  -.0149061   .0608735    -0.24   0.807     -.135373    .1055608
        eap9 |   .0289287   .0581808     0.50   0.620    -.0862095    .1440668
       eap10 |   .1221601   .0745059     1.64   0.104    -.0252849    .2696051
       eap11 |  -.0936096   .0399802    -2.34   0.021    -.1727293     -.01449
       eap12 |  -.0206429   .0429898    -0.48   0.632    -.1057185    .0644327
  _Iperiod_2 |   .1150264    .031423     3.66   0.000     .0528412    .1772116
  _Iperiod_3 |   .0813417   .0261022     3.12   0.002     .0296862    .1329972
  _Iperiod_4 |    .127689    .026469     4.82   0.000     .0753077    .1800703
  _Iperiod_5 |   .1340237   .0300877     4.45   0.000     .0744809    .1935664
  _Iperiod_6 |   .0463128   .0224257     2.07   0.041     .0019329    .0906926
  _Iperiod_7 |   .0283441   .0237049     1.20   0.234    -.0185672    .0752553
  _Iperiod_8 |  -.0578551   .0212537    -2.72   0.007    -.0999155   -.0157948
  _Iperiod_9 |  -.0187763   .0210113    -0.89   0.373     -.060357    .0228045
 _Iperiod_10 |  -.0823159   .0261085    -3.15   0.002    -.1339839   -.0306479
 _Iperiod_11 |   .0213753   .0176483     1.21   0.228    -.0135502    .0563009
```

```
 _Iperiod_12 |   .0298809   .0154189     1.94   0.055    -.0006326    .0603945
       _cons |   .2696911   .1541875     1.75   0.083    -.0354414    .5748236
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ea eap2 eap3 eap4 eap5 eap6 eap7 eap8 eap9 eap10 eap11 eap12
    eap13 _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7
    _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.29  Pr > z =  0.022
Arellano-Bond test for AR(2) in first differences: z =  -0.77  Pr > z =  0.440
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 317.55  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  = 107.09  Prob > chi2 =  0.452
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)    =  77.19  Prob > chi2 =  0.185
    Difference (null H = exogenous): chi2(39)    =  29.90  Prob > chi2 =  0.852
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)    =  40.40  Prob > chi2 =  0.036
    Difference (null H = exogenous): chi2(80)    =  66.70  Prob > chi2 =  0.856
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)    =  98.35  Prob > chi2 =  0.043
    Difference (null H = exogenous): chi2(30)    =   8.74  Prob > chi2 =  1.000
  iv(ea eap2 eap3 eap4 eap5 eap6 eap7 eap8 eap9 eap10 eap11 eap12 eap13
_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7 _Iperiod_8
_Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13, eq(level))
    Hansen test excluding group:     chi2(84)    =  98.74  Prob > chi2 =  0.130
    Difference (null H = exogenous): chi2(22)    =   8.35  Prob > chi2 =  0.996

//The HCASM with SSA, EA, SSAP and EAP
Note: The number of instruments is reported in the STATA output is the total
number of instruments used for all endogenous, predetermined and exogenous
variables.
i.period          _Iperiod_1-13       (naturally coded; _Iperiod_1 omitted)
Favoring speed over space. To switch, type or click on mata: mata set matafavor
space, perm.
ssap2 dropped due to collinearity
ssap3 dropped due to collinearity
ssap13 dropped due to collinearity
eap2 dropped due to collinearity
eap13 dropped due to collinearity
_Iperiod_13 dropped due to collinearity
Warning: Number of instruments may be large relative to number of observations.
Warning: Two-step estimated covariance matrix of moments is singular.
  Using a generalized inverse to calculate optimal weighting matrix for two-
step estimation.
  Difference-in-Sargan/Hansen statistics may be negative.
Dynamic panel-data estimation, two-step system GMM
------------------------------------------------------------------------------
Group variable: id                              Number of obs      =      1114
Time variable : period                          Number of groups   =       127
Number of instruments = 143                     Obs per group: min =         3
```

```
F(36, 126)    =   1096.59                                      avg =      8.77
Prob > F      =     0.000                                      max =        12
------------------------------------------------------------------------------
             |              Corrected
        ln_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_y |
         L1. |   .9690403   .0167629    57.81   0.000     .9358671    1.002214
             |
        ln_s |   .0652122   .0312385     2.09   0.039     .0033922    .1270322
      ln_ngd |   .0538383     .06605     0.82   0.417    -.0768728    .1845494
        ln_h |   .0449139   .0392978     1.14   0.255    -.0328552     .122683
         ssa |  -.0133551   .0331706    -0.40   0.688    -.0789988    .0522885
          ea |   .0616782   .0363304     1.70   0.092    -.0102185    .1335749
       ssap4 |  -.0573072   .0421633    -1.36   0.177    -.1407471    .0261326
       ssap5 |  -.0507504   .0492587    -1.03   0.305    -.1482319    .0467311
       ssap6 |  -.0535255   .0559607    -0.96   0.341    -.1642701     .057219
       ssap7 |  -.1145114   .0616362    -1.86   0.066    -.2364876    .0074648
       ssap8 |  -.0498116   .0552212    -0.90   0.369    -.1590927    .0594695
       ssap9 |  -.0359428   .0470159    -0.76   0.446    -.1289859    .0571002
      ssap10 |  -.1135311   .0494936    -2.29   0.023    -.2114775   -.0155848
      ssap11 |  -.0902881   .0462503    -1.95   0.053    -.1818161    .0012399
      ssap12 |  -.0922159   .0333561    -2.76   0.007    -.1582266   -.0262052
        eap3 |  -.1719074   .2363484    -0.73   0.468     -.639634    .2958191
        eap4 |  -.1626053   .0645216    -2.52   0.013    -.2902916    -.034919
        eap5 |   -.071205   .0970369    -0.73   0.464    -.2632381    .1208281
        eap6 |  -.0519266   .0757024    -0.69   0.494    -.2017393    .0978862
        eap7 |  -.0160553   .0780736    -0.21   0.837    -.1705606    .1384501
        eap8 |  -.0380797   .0544728    -0.70   0.486    -.1458797    .0697203
        eap9 |   .0133258   .0556043     0.24   0.811    -.0967135    .1233651
       eap10 |   .0909987   .0708278     1.28   0.201    -.0491674    .2311649
       eap11 |  -.1130031   .0376194    -3.00   0.003    -.1874508   -.0385555
       eap12 |  -.0339665   .0462816    -0.73   0.464    -.1255565    .0576235
  _Iperiod_2 |   .1336009   .0344201     3.88   0.000     .0654845    .2017173
  _Iperiod_3 |   .0967678    .027405     3.53   0.001      .042534    .1510015
  _Iperiod_4 |   .1422175   .0265101     5.36   0.000     .0897548    .1946801
  _Iperiod_5 |   .1453469   .0299961     4.85   0.000     .0859855    .2047082
  _Iperiod_6 |    .056788   .0230161     2.47   0.015     .0112398    .1023363
  _Iperiod_7 |   .0486031   .0253591     1.92   0.058    -.0015819     .098788
  _Iperiod_8 |  -.0370428   .0264485    -1.40   0.164    -.0893835     .015298
  _Iperiod_9 |  -.0089459   .0266717    -0.34   0.738    -.0617284    .0438365
 _Iperiod_10 |  -.0398623   .0306305    -1.30   0.195    -.1004792    .0207546
 _Iperiod_11 |   .0381937   .0175305     2.18   0.031     .0035014    .0728859
 _Iperiod_12 |   .0499528   .0186786     2.67   0.008     .0129883    .0869173
       _cons |   .5439207   .2156857     2.52   0.013      .117085    .9707563
------------------------------------------------------------------------------
Instruments for first differences equation
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    L(2/3).ln_h
    L(3/4).(ln_y ln_s ln_ngd)
Instruments for levels equation
  Standard
    ssa ea ssap2 ssap3 ssap4 ssap5 ssap6 ssap7 ssap8 ssap9 ssap10
    ssap11 ssap12 ssap13 eap2 eap3 eap4 eap5 eap6 eap7 eap8 eap9
    eap10 eap11 eap12 eap13 _Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5
    _Iperiod_6 _Iperiod_7 _Iperiod_8 _Iperiod_9 _Iperiod_10 _Iperiod_11
    _Iperiod_12 _Iperiod_13
    _cons
  GMM-type (missing=0, separate instruments for each period unless collapsed)
    DL.ln_h
    DL2.(ln_y ln_s ln_ngd)
------------------------------------------------------------------------------
Arellano-Bond test for AR(1) in first differences: z =  -2.44  Pr > z =  0.015
```

```
Arellano-Bond test for AR(2) in first differences: z =  -0.75  Pr > z =  0.451
------------------------------------------------------------------------------
Sargan test of overid. restrictions: chi2(106)  = 305.10  Prob > chi2 =  0.000
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(106)  =  97.28  Prob > chi2 =  0.716
  (Robust, but weakened by many instruments.)
Difference-in-Hansen tests of exogeneity of instrument subsets:
  GMM instruments for levels
    Hansen test excluding group:     chi2(67)   =  73.05  Prob > chi2 =  0.286
    Difference (null H = exogenous): chi2(39)   =  24.24  Prob > chi2 =  0.969
  gmm(ln_y ln_s ln_ngd, lag(3 4))
    Hansen test excluding group:     chi2(26)   =  29.27  Prob > chi2 =  0.299
    Difference (null H = exogenous): chi2(80)   =  68.01  Prob > chi2 =  0.828
  gmm(ln_h, lag(2 3))
    Hansen test excluding group:     chi2(76)   =  87.18  Prob > chi2 =  0.179
    Difference (null H = exogenous): chi2(30)   =  10.10  Prob > chi2 =  1.000
  iv(ssa ea ssap2 ssap3 ssap4 ssap5 ssap6 ssap7 ssap8 ssap9 ssap10 ssap11
ssap12 ssap13 eap2 eap3 eap4 eap5 eap6 eap7 eap8 eap9 eap10 eap11 eap12 eap13
_Iperiod_2 _Iperiod_3 _Iperiod_4 _Iperiod_5 _Iperiod_6 _Iperiod_7 _Iperiod_8
_Iperiod_9 _Iperiod_10 _Iperiod_11 _Iperiod_12 _Iperiod_13, eq(level))
    Hansen test excluding group:     chi2(74)   =  82.54  Prob > chi2 =  0.232
    Difference (null H = exogenous): chi2(32)   =  14.74  Prob > chi2 =  0.996
```